NTNU
Norwegian University of
Science and Technology

Pervasively Gamifying the Museum Experience

An empirical investigation of knowledge gain
and engagement

Jamawadi Noor

Master of Science in Informatics
Submission date: December 2016
Supervisor: Maria Letizia Jaccheri, IDI
Co-supervisor: Michail Giannakos, IDI

Norwegian University of Science and Technology
Department of Computer Science

Summary

Museum experiences are naturally one-shot experiences with low revisit rates and cater to sprawling classrooms where teachers often have reduced oversight. At times the open-endedness of museums distracts museum-goers. This is an area that gamification provides great motivational benefits that might reduce these problems, in addition to digital systems often enabling a higher degree of oversight and one-to-one feedback. Pupils and other museum-goers alike benefit from increases to engagement and knowledge gain. Pervasive gamification presents itself as an opportunity in order to retain the open-endedness and exploratory nature of museums. Pervasive gamification also solves many of the oversight issues teachers are plagued with, as each user retains a profile seamlessly across the many different clients that enables assessment and increased sense of control.

The main goal of this research was to investigate whether a pervasively gamified museum experience can provide knowledge gain and user engagement benefits. In the process of answering these questions, we developed an application — peMuse — that employs a client-server architecture with seamless RFID-based authentication for pervasiveness. The application was built with a foundation of gamification techniques and teaching methods inspired by behavioural psychology. After concluding the development, we conducted a field-study with 36 participants over two days. Our research concludes with two slightly contrasting results. Our application scores high in user-engagement, yet shows no statistically significant increase in knowledge gain. We identified possible causes for the lack significant knowledge gain. Qualitative results gathered also attest to an overwhelmingly positive feedback, with participants experiencing increased engagement.

Sammendrag

Museumsbesøk er et typisk engangstilfelle — de fleste besøker museer en gang eller mindre i løpet av et år. Museumsbesøk er for lærere ofte uoversiktlig, og grunnet mindre en-til-en kommunikasjon kan enkelte elever få lite utbytte. Av og til kan mangel på rammer begrense lysten og motivasjonen til besøkende til å utforske alt som er tilgjengelig. Dette er et område hvor gamification (spillifisering) tilbyr store fordeler innenfor motivasjon som kan sammen med det digitale aspektet tilby høyere nivåer med oversikt og mer individuell tilbakemelding. Både elever og andre museumsbesøkende vil oppleve økende engasjement og kunnskapsøkning. Pervasive gamification er en mulig løsning som både ivaretar åpenheten og den utforsknende delen av det å gå på museum. Pervasive gamification løser også mange av oversiktsproblemene som lærere kan oppleve fordi hver bruker har sin egen profil som blir sømløst ivaretatt fra klient til klient.

Hovedmålet for denne forskningen er å utforske om pervasive gamification av et museumsbesøk kan tilby kunnskapsøkning og økt engasjement. For å svare på disse spørsmålene har vi utviklet en applikasjon — peMuse — som har en klient-server arkitektur med sømløs RFID autentisering. Applikasjonen er bygget på gamification teknikker og læringsmetoder inspirert av atferdspsykologi. Etter fullført utvikling, gjennomførte vi et eksperiment med 36 deltakere. Resultatet var økt engasjement, men ikke statistisk signifikant økning av kunnskap.

Acknowledgements

This project was inspired by a previous application developed by Ioannis Leftheriotis, his application provided a foundation for early prototyping and design, and the research bootstrapped much of our own literature review. I would also like to acknowledge my supervisor Michail Giannakos for great feedback and advice in research areas to focus on, as he also had a hand in the previous application. The experiments conducted were worked out with the help of Kshitij Sharma, Ioannis Leftheriotis and Michail Giannakos. I would also like to acknowledge Kshitij Sharma for domain choice and help designing the posters, and setup of the proxy-museum. Extra acknowledgements are extended to Francesco Gianni for valuable and frequent aid with poster printings and Tord Kloster for early prototyping user-tests.

Table of Contents

List of Tables

List of Figures

Abbreviations

MT	=	Multi-Touch
GAID	=	Gamified Interactive Display
PICC	=	Proximity Integrated Circuit Card
PCD	=	Proximity Coupling Device
RFID	=	Radio-frequency Identification
GAID	=	Gamified Interactive Display
CSCL	=	Computer-Supported Collaborative Learning
UI	=	User Interface

Chapter 1

Introduction

Since 2010, gamification has gone from debuting on Google (Google Trends, 2016) to having an estimated market value of $1.65 billion in 2015 and is estimated to reach a predicted market value of $11.10 billion by 2020 (P&S Market Research, 2016). We define gamification according to Deterding et al. (2011): "Based on our research, we propose a definition of "gamification" as the use of game design elements in non-game contexts". In this thesis we attempt translate gamification techniques in the context of the museum.

Museums are academic institutions whose primary roles are the preservation, articulation and exhibition of artifacts and should serve as places of education and entertainment Prince (1990). Exploring how the educational and entertainment aspects of museums can be improved is the goal of this thesis, and the intention is to achieve this via technological artifacts such as multi-touch (MT) screens, well researched and implemented gamification techniques, and a foundation of informal computer-supported collaborative learning (CSCL) research.

The work presented extends upon previous work on gamified interactive displays (GAID) presented by my supervisor. More information on this original application can be found in chapter 3.1.2. A prototype pervasive system was produced named peMuse (a semi-portmanteau of pervasive and museum), inspired by the original application but with significant feature additions. There is no longer just the application, but rather a system, including RFID sensors and chips, multiple Arduinos, multiple MT devices, and a RESTful client-server architecture working in unison. In the next section we outline our motivation for conducting this research.

1.1 Motivation

Museum experiences are one-shot in nature as museums typically don't see the same visitors more than a few times a year. According to a large 2004 survey conducted by The Government of Canada (2004), 39% of responders stated that they visit museums of any size more than once a year. The rest (61%) were at once a year (608,725) or less (372.765). This fell in line with personal hypotheses regarding re-visitation rates. In addition to low repeat visitations, school-groups — one of the groups museums cater most to — frequently suffer from lowered oversight, lack of student motivation and limited teacher-student interaction compared to classrooms DomíNguez et al. (2013). Online learning helped solve some of the latter issues in other out-of-school contexts, how can we apply this to actual museums? Here lies our motivation, the issues presented with regards to visitation rates and school-groups can be solved using gamification. Visitors can be more engaged and satisfied, and gamifying the museum experience would provide similar benefits to classrooms due to reduced oversight for teachers whilst in unfamiliar and asymmetric environments.

Keeping in mind the one-shot interactions and the demographics we are focusing on, a more gamified approach to this informal learning setting could yield significant benefits in user engagement and learning. A generic method of accomplishing this poses significant benefits to society.

1.2 Research Questions

We have two research questions:

- RQ1: Does a pervasively gamified museum experience increase knowledge gain?
- RQ2: Does a pervasively gamified museum experience increase user engagement?

1.3 Research strategies

This thesis has two research strategies:

- Design and creation: The design and creation (Oates, 2005) of a gamified interactive display instantiation (i.e. working system, prototype) with pervasive elements.
- Experiments: Executing a field study using the previously created application.

The contributions are therefore twofold; the application itself as a prototype and proof of concept for the generic pervasive gamification, and the results and analyses of the field studies conducted. Oates (2005) suggests that "Typically one research question has one research strategy". We don't intend to assign each research question to each strategy here; both research questions are to be answered via the experiments. However this necessitates the existence of the artifact which the first research strategy produces.

1.3.1 Design and creation

This research strategy involves the design and creation of the system of IT artifacts that serves as a vehicle for the subsequent experiments. This process will itself be regarded a part of the research and part of the knowledge contribution this thesis represents.

We use the five-step design and creation process (Vaishnavi & Kuechler, 2015), see table 1.1

We will examine what happens when this artifact is subjected to real testing in what approximates a real-life context, the results gathered during the field study will inform a discussion and conclusion.

Table 1.1: Design and creation process

Name	Description
Awareness	We will be studying the state-of-the-art, collecting authors' suggested areas of further research and using this to articulate a concise problem
Suggestion	We will suggest a solution to this aforementioned problem
Development	We will iteratively design and develop the system based on the suggestion
Evaluation	We will examine the developed artifact, assess to what degree the artifact fulfills our suggestion. Any deviations from expectations are noted and explained before moving on to the experiments
Conclusion	As this artifact only serves as a vehicle for the field study, the conclusion will be shared within this thesis

1.4 Thesis Structure

- Chapter 2 - An overview of definitions, state-of-the-art related work on gamification, museums, serious games, education etc.
- Chapter 3 - Designing and implementing a solution that satisfies our problem
- Chapter 4 - Research questions and methodology outlining our research design and experiments
- Chapter 5 - Results from the experiment
- Chapter 6 - Discussion
- Chapter 7 - Conclusion

Chapter 2

Background

There is a significant number of fields we are delving into; educational games, serious games, MT technology, gamification, pervasive games, CSCL and museums. In this chapter we establish definitions and review previous research relating to above keywords in a satisficing manner.

2.1 Definitions

2.1.1 Gamification

As mentioned earlier, gamification is a relatively young field, the term only starting to trend in 2010 (Google Trends, 2016) (see fig 2.1), and since then grown into a multi-billion dollar market (P&S Market Research, 2016). The rapid adoption of game-like elements into a wide variety of software begged for an examination of whether this truly was a new phenomenon or a re-branding of existing techniques. Deterding et al. (2011) argue that gamification is a new and distinct enough phenomena and propose the following definition: "“Gamification” is the use of game design elements in non-game contexts". Huotari & Hamari (2012) argue that Deterding's aforementioned "game elements" is too vague as there is no clearly defined set of game elements unique to games. They argue that gamification is to be understood as a process in which a "gamifier is attempting to increase the likelihood for the gameful experiences to emerge by imbuing the service with affordances for that purpose". Of note is Houtari & Hamari's highlighting of gameful experiences being voluntary and led by users' intrinsic motivations. We believe Houtari & Hamari bring valuable nuance with their definition, however we will be using the definition provided by Deterding et al. as it succinctly encapsulates the core of gamification.

2.1.2 Serious Games

Deterding et al. (2011) also delve into defining serious games, arguing that serious games "... describes the design of full-fledged games for non-entertainment purposes". But also

Figure 2.1: The birth of gamification.

noting the fuzzy and subjective border between games and game-like artifacts. We define serious games according to Susi et al. (2007), as games that are used for more than entertainment purposes only. We state that peMuse — our application — is not a game, though it does adopt game-like elements. The context that is being gamified pervades the computer software itself, i.e. the museum context is our non-game context as described by Deterding et al. (2011), which leads us to the next subsection.

2.1.3 Pervasive Games

Pervasive games — or pervasive computing — integrate the physical with the digital and bring the gameplay into the physical world (Benford et al., 2005). In the context of this thesis it is beneficial to define how pervasive computing integrates with our system. The multiple MT interfaces, Proximity Integrated Circuit Cards (PICCs) that players carry as authentication, and the change of content dependent on location together make clear that the boundaries of the screen does not contain the game. Users compete and collaborate with one another keeping in mind their scores and available power-ups. Users are able to engage with the real-world in order to improve their scores. The non-game context being gamified is being gamified in a pervasive manner, which leads us to our own holistic term; pervasive gamification - the use of game design elements in a non-game physical-world context.

2.2 Related work

In this subsection we review articles, books and papers accumulated during the literature review that were found to be relevant to the keywords; educational games, serious games, MT technology, gamification, pervasive games, CSCL and museums.

2.2.1 Learning through Multi-touch Interfaces in Museum Exhibits: An Empirical Investigation

Figure 2.2: “Walls of Nicosia” running on a MT table

“Walls of Nicosia”, a 3D MT application was used as the artifact in this empirical investigation. The interactive application allows users to interact with a 3D environment using MT as a virtual tour of the different walls of Nicosia throughout history. The goal of the application is to “explain and present the development of the area and the history of the development of the fortifications of the city” (Zaharias et al., 2013). There was a scientific control experiment conducted; a control group (referred to as traditional), and an experiment group (referred to as virtual). Participants were randomly selected from elementary schools in Nicosia, 24 girls and 29 boys aged 10 to 11 years old. Data collection was conducted using both pre- and post-test questionnaires. Data analysis of the results revealed there was no statistically significant difference in learning performance between the virtual and traditional group. However user-experience scores were higher in the virtual group, users even expressing intentions to repeat visits.

Zaharias et al. (2013) argue that the novelty of the application might have distracted the students, negatively affecting their learning performance. Due to the one-shot nature of museums this issue may not be resolvable. An argument could be made that museums - as technologically advanced as they are - offer such novelties regardless. There is no mention of how the user-experience scores being higher might have a causative link with the novelty factor. They also mention that “...after the novelty effect has passed, such interactive technologies can provide more authentic learning and entertainment at the same time” (Zaharias et al., 2013). I disagree with this conclusive hypothesis once again due to the one-shot nature of museum visits, there is obviously no reason that visitors cannot visit a museum more than once. I would argue though that the novelty of such installations in museums only passes after a few closely timed visits, which may turn out to be improbable.

The limitations of the study discussed here relate to length of knowledge tests, the focus on short-term retention (similar to skill and drill), lack of avatar and difference in teachers for the two groups and their different influences.

2.2.2 Designing playful games and applications to support science centers learning activities

This paper (Giannakos et al., 2014) presents a set of design considerations that are very relevant for our system. The intent of the research was to investigate how games and applications could be designed to motivate students for informal learning situations. The following is an abridged list of these guidelines:

1. The structure must be easy and understandable
2. Use visual and interactive elements
3. Incorporate entertaining responses for On-Screen actions
4. Enhance interaction with large touch-screens
5. Apply short-term scalable design (Include multiple short stages)
6. Direct users' attention to specific topics
7. Support cooperation and competition among users
8. Connect the activity with the school curriculum emphasize

Although these guidelines are created with STEM fields in mind, we believe this is an appropriate foundation for our design choices given our goal of increasing both knowledge gain and user engagement.

2.2.3 Practical, appropriate, empirically-validated guidelines for designing education game

This another paper outlining design guidelines for educational games (Linehan et al., 2011). These guidelines are developed with a background of validated teaching methods developed by behavioural psychologists and presents a practical framework for designing engaging game mechanics. The core of this framework is Applied Behaviour Analysis (ABA), which is a practical teaching method that involves modifying behaviour via learning principles. In our case the goal is to increase knowledge gain and engagement, as such the behaviour we want to promote is learning and motivation.

Although we are explicitly not designing a game, we are introducing game-like elements in a gamified manner. Linehan et al. (2011) describe how analysis and information gathered can be useful tool for regulating rewards for maximum effect with regards to behavioural goals (i.e measure impact of rewards, provide more of the reward that user seems to be motivated to gain). The attention to ABA is perhaps the most interesting revelation from this paper, and we intend to incorporate this into our application.

2.2.4 Gamification in theory and action: A survey

This paper presents a survey of gamification uses in published theoretical reviews and research papers. It contains within it systematic reviews of gamification in action, design, approaches and techniques, and user impact. The findings related to the user impact effectiveness of gamification is "…positive-leaning but mixed…" (Seaborn & Fels, 2015). This aligns with the findings of DomíNguez et al. (2013), although game-design choices bear some of the responsibility in that case. Of special interest is the suggestions from Deci et al. (1999) research that gamification effects are temporary or damaging over time, this drawback will most likely not apply to our end-users exposure to our solution is limited. Seaborn & Fels (2015) also note: "The present body of applied gamification research suggests that success might be improved across the board if the design of gamified systems – especially extrinsic motivators – is informed by end-users' intrinsic motivators".

One of the conclusions is that gamified systems do not have an optimal combination of game-design elements, but instead should rely on user-centred design approaches to select and configure these elements for a given end-user population (or customisation options).

2.2.5 BATs and APEs: Designing an interactive tabletop game for natural history museums

Horn et al. (2012) describe visitor interactions with an application developed for natural history museums using interactive tabletops. The goal of the research was to investigate successful approaches to designing interactive tabletop exhibits for museums. Though this application is itself an exhibit and our large MT application is intended to be placed adjacent to an exhibit, its findings - especially with regards to informal CSCL - are relevant for our research. The game, BAT (Build-a-Tree), was featured in a field study with 35 groups participating and a further 50 social groups with lessened observation.

The research conclusions highlighted advantages of tabletop games for learning in museums, citing motivational benefits as well as social game-play practices users inherently employ that aids collaboration. As important are the pitfalls observed by the authors, such as multilevel games with increasing difficulty that have groups wander in mid-game. This leads to confusion and frustration for subsequent groups as the earlier levels scaffold game-play. Solutions to this involve making it easier for users to restart the application, decreasing the amount of time required for the game, and implementing a time-out feature. The second pitfall was visitors aversion to playing games, there are many solutions to this and the user's own attitude plays a large role. The authors suggest placing artifacts adjacent to the table-top in order to associate the game with the museum's content. A final caution relayed by the authors identifies a pedagogical trade-off that might occur when designing such games; will relying on games such as BAT be at the expense of exploratory and open-ended activities that normally take place in museums? We will attempt to address this issue in our design by not enforcing any specific order to the MT devices, allowing for similar open-ended exploration. In conclusion, tabletops were found to be a compelling device for walk-up-and-use form that supports CSCL.

2.2.6 Deconstructing the touch experience

Although the benefits on MT surfaces are perhaps intuitively understood as superior to the alternative (multiple mice) in the context of our proposed application, we still find it pertinent to understand and explain why. This paper (Watson et al., 2013) evaluates performance and experience between touch and mouse input devices on both horizontal and vertical surfaces. A study was conducted with 48 participants. The results concluded no effect on orientation (horizontal/vertical) and touch outperforming mouse input in both speed and accuracy. The authors also conclude that touch scored higher than mouse input in participant happiness, enjoyment, competence, control, relatedness, and immersion. Watson et al. also describe how one can design for the touch experience, stating that "...feelings of competence through touch can be used when designing interfaces for novices...'. We aim to support the benefits that touch interfaces supplies with a large MT display.

2.2.7 The Hook model

A crucial requirement in making gamification work for our system is a framework of motivational science. The hook model (Eyal, 2014) attempts to provide just that in the form of an interaction loop. In the gaming and marketing industry there is a design pattern used to build habit-forming products, Eyal (2014) describes it as a "hook". This design pattern describes four fundamental stages of a recurring user interaction loop (fig 2.3):

1. Trigger
2. Action
3. Reward
4. Investment

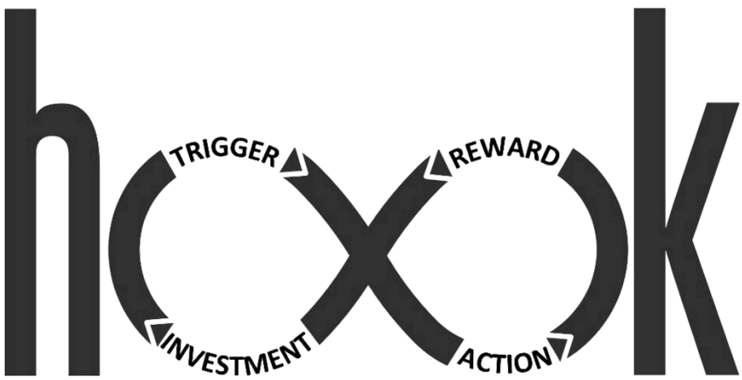


Figure 2.3: The hook model

The hook model will describe the base of our interaction loop. In short, the initial trigger (1) is the introduction of a PICC to a user, who then performs the action (2) of solving the puzzle on the terminal; the GAID. Upon completion they are rewarded (3) and hopefully the rewards instill a sense of investment (4). They are then prompted to seek out the next terminal and the loop continues. When the last terminal is completed users have their accomplishments displayed. Another benefit of the PICC badge is the ability to make any terminal the last terminal (by counting number of completed puzzles).

As Seaborn & Fels (2015) suggest, success when employing gamification might be improved by informing design decisions on end-users intrinsic motivations. Eyal (2014) refers to these as internal triggers: intrinsic motivations that manifest in the minds of the users. An external trigger is referred to as a sensory stimuli, e.g. a phone ringing. The interaction starts with triggers, external and internal, that prompt us to begin an action. Eyal defines these actions as the simplest behaviour done in anticipation of a reward. This is inline with the Fogg (2002) behaviour model (FBM). Using FBM we can predict that Behaviour = Motivation * Ability * Trigger. This behaviour model asserts that "for a target behavior to happen, a person must have sufficient motivation, sufficient ability, and an effective trigger". In our case the intended behaviour is to play and complete the puzzles at the client. Assuming the puzzles are well designed in the "ability" capacity (i.e. users have sufficient ability to complete them), the remaining obstacle is the user's motivation, how do we increase the user's motivation? The next subsection features a gamification design framework that we can use to answer this question.

2.2.8 The Octalysis framework

Seaborn & Fels (2015) concluded that there is no optimal configuration of game-elements for gamification, but that user-centred design or high customisability would lead to a best fit. By using a gamification design framework we can map gamification techniques to our targeted users. Octalysis (Chou, 2013) is a gamification design framework that maps out 8 core drives behind human motivation. Fig 2.4 illustrates the framework, it is divided into four overlapping segments (not illustrated). The top three — Meaning, Accomplishment and Empowerment — are called white hat gamification techniques, while the bottom three — Avoidance, Scarcity and Unpredictability — are black hat. Informally black hat techniques exploit human behaviour whilst white hat empowers them. The leftmost and rightmost three both describe different types of core drives, respectively; extrinsic and intrinsic. Extrinsic drives are ignited by a goal, reward or achievement, such as leveling up, experience points, trophies etc. Intrinsic drives come from within, the experience is enjoyable in and of itself, e.g. curiosity, need for positive feedback and creativity empowerment.

Octalysis serves as a great framework to analyse which gamification techniques fits best for our solution. We intend to use this framework to holistically quantify the core drives behind our solution for analysis purposes. This list of drives to focus will We also intend to use this framework to give insight into our gamification technique choices below.

2.2.8.1 Drives to focus on

The one-shot nature of the museum experience helps narrow down our choices for us, whatever accumulated rewards and achievements the user has does not need to be stored for the next time they visit. In descending order of importance, the core drives we should focus on are:

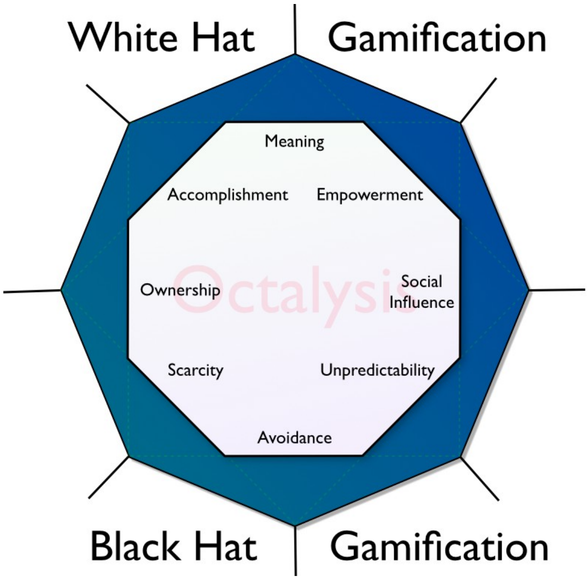


Figure 2.4: The Octalysis framework

1. Empowerment of Creativity & Feedback
2. Social Influence & Relatedness
3. Unpredictability & Curiosity
4. Meaning & Calling
5. Development & Accomplishment
6. Ownership & Possession
7. Loss & Avoidance
8. Scarcity & Impatience

We rely less on black hat techniques and more on white hat techniques with the intention of leaving a sense of achievement and empowerment in the user. Intrinsic core drives (unpredictability, social influence and empowerment) are also preferred compared to the extrinsic ones (Seaborn & Fels, 2015). Hopefully focusing on these techniques will help motivate users from within, fostering an internal desire to complete puzzles and learning, and impact cognitive, emotional, and social areas of users (Lee & Hammer, 2011). It is important to remember the focus on education when considering gamification techniques, for instance relying heavily on rewards handicaps the user's willingness to take risks, which in turn restricts creativity (Amabile et al., 1990). Our design should encourage experimentation.

2.2.8.2 Gamification feature list

This list represents the output of an early Octalysis-based brainstorming session with the related work and literature in mind. Brainstorming occurred during several sessions, keeping the target user demographic — pupils and classrooms — in mind.

Figure 2.5: Octalysis brainstorming

Note that table 2.1 isn't a final feature list, but the result of the first brainstorming session with the related work in mind. With this foundation we start iterating on prototypes.

2.3 Summary

In this chapter we've established crucial definitions before moving on to the state-of-the-art. We've gathered several papers and collected key guidelines (Giannakos et al., 2014; Watson et al., 2013), cautions of pitfalls (Horn et al., 2012), and discovered potential improvement points (Zaharias et al., 2013) that we can apply to our designs. We've also focused on behavioural analysis (Linehan et al., 2011; Eyal, 2014) and taken advice of Seaborn & Fels (2015) to employ frameworks for working out the optimal combination of game-design elements (Chou, 2013). Our design will follow well researched guidelines, improving on shortcomings, avoiding pitfalls, and having design informed by behavioural psychology, and include one point of innovation; pervasiveness. As we defined pervasiveness earlier (Benford et al., 2005), our design will attempt to solve issues posed by Horn et al. (2012) on lack of open-endedness and exploration, and challenge the call for exploring new contexts to gamify (Seaborn & Fels, 2015). Our pervasive approach — with the aforementioned state-of-the-art as a foundation — will form the substance of this thesis.

Table 2.1: Brainstormed gamification features

Feature	Octalysis technique
Terminal shall have a prompt as external trigger for the start-up of a game.	Meaning and calling
Users will receive a number of experience points (XP) upon puzzle completion.	Empowerment of Creativity, Development
For every 1000 XP the user receives an achievement.	
The number of points received does not decrease if the user has previously failed.	
User shall always have XP displayed.	Development, Feedback
Upon successful puzzle completion the user will receive feedback in the form of sound, text and animation.	
Upon hidden leveling the avatar should reflect the current level with new graphics	
Variable reward chance of receiving a temporary boost in the form of an icon once touched gives the user a temporary boost. The probability increases if the user is using more time than normal on puzzles. These icons are always present and transparent when you have none, they also stack.	Empowerment of Creativity, Unpredictability, Ownership
Freeze time - Stop timer for 1 minute	
Cash in - Receive 100xp immediately	
Hint? - A hint pertaining to the question is displayed until closed	
Users should be able to collaborate with one other user using the same GAID, using this will introduce a flat collaboration bonus XP.	Social Influence
Once the final puzzle/terminal is completed the user is asked to enter a name, the user will then have their score displayed along with recent players.	
The display shall resemble a scrollable table with each row displaying respective points, avatar, color, name, level and achievements sorted by points.	

Chapter 3

System Design and Implementation

In this chapter we discuss how we iterated through several designs, and how the final solution was implemented.

3.1 Design

We refer once again to the design and creation process by Vaishnavi & Kuechler (2015). The design portion of this process involves (1) awareness, and (2) suggestion (design and creation process table). In the previous chapter we collected and discussed related work and the state-of-the-art. In this section we will establish the gap, or improvement points that our artifact will attempt to incorporate.

One of the important factors of this thesis contributions is innovation. We move towards the edge of knowledge by exploring the space (Seaborn & Fels, 2015), quite literally at that as we attempt to pervasively gamify the museum space. The game should extend into the museum space and track individual users profiles, for this reason we decide to use an analogue to RFID attendance systems (Chew et al., 2015). This is also inspired by the Eyal (2014) take on triggers — in particular external ones — as the PICC (i.e. the RFID tag) will in itself as a physical device (that users receive on entry) function as an external trigger. Zaharias et al. (2013) suggestion to use avatars also functions well with this system, as the PICC's unique identification number (UID) can link to a certain avatar that users retain throughout their visit. This meld of quick authentication, tracking of users as they move throughout the space in an open-ended explorative manner (Horn et al., 2012) creates the baseline of what we call a pervasively gamified museum experience. The following subsection lists our requirements.

3.1.1 Requirements

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 (Bradner, 1997). This requirement specification is from during the design phase. All deviations from this requirement list during development are mentioned in the chapter 3.2.

With these requirements we can work on designing prototypes in order to see how it would look. The utility of this application is wholly dependent on usability, so great care was taken to develop with a UX-first approach (Abras et al., 2004).

3.1.1.1 Must

The application must:

1. Be able to use a provided PICC with client's PCD to authenticate and fetch valid profile and avatar.
2. Provide a single unique consistent avatar for each user across multiple clients until a manual reset is done.
3. Show a start-screen upon authentication for each player, screen split into columns equalling number of players.
4. Be able to play with another player in one of two modes; versus or co-op.
5. Have a versus mode where only correct players are rewarded with xp.
6. Have a co-op mode where all players are rewarded if the correct choice is made.
7. Have a co-op mode with one set of interactive choice buttons.
8. Provide external trigger that prompts users to engage with client.
9. Must reward users with a predetermined amount of xp immediately after each correct answer.
10. When user is rewarded that user's profile should be updated across all devices.

3.1.1.2 Should

The application should:

1. Be multi touch capable; able to handle 10 points simultaneously.
2. Be able to retain accrued points and achievements at any terminal using the provided PICC to authenticate.
3. Always show the user their current progress in terms of xp and level
4. Be a variable chance of powerup gain after successful answer
5. Be at least two different power-ups that provide a tangible benefit to user that uses it in that round.

3.1.1.3 May

The application may:

1. Have an administration panel for changing questions.
2. Have support for more than two players.
3. Display high-score list

3.1.2 Original Application

This master thesis started out with the task of improving a previously existing Gamified Interactive Display (GAID) application. There were two examples, an application for a library-museum and one for a deaf-museum. The applications' design was grounded on three epistemological traditions: (1) behaviourism (knowledge is gained through experience), (2) cognitivism (knowledge is negotiated through experience and thinking) and (3) constructivism (knowledge is constructed). For this reason the application had three different stages linked to the three traditions:

1. Exploration stage (fig 3.1)
2. Application stage (fig 3.2)
3. Creation stage (fig 3.3)

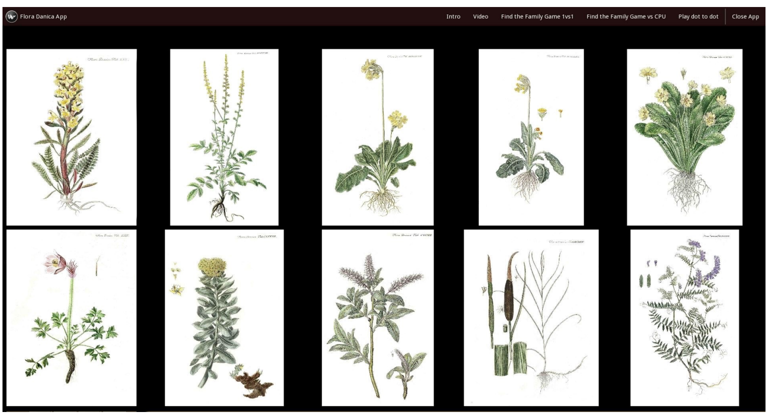


Figure 3.1: Deaf-museum: Exploration stage

As the figures (3.1, 3.2, 3.3) depict, there is a tight coupling between base theory and final artifact. Given the new goal of pervasiveness for the application we believe that museum exhibits themselves should function as the exploration stage, opting to leave this out of the application itself. This requires the GAID to be placed within a proper vicinity of the exhibits obviously. The application stage will remain the same, a quiz that can be

Figure 3.2: Deaf-museum: Application stage

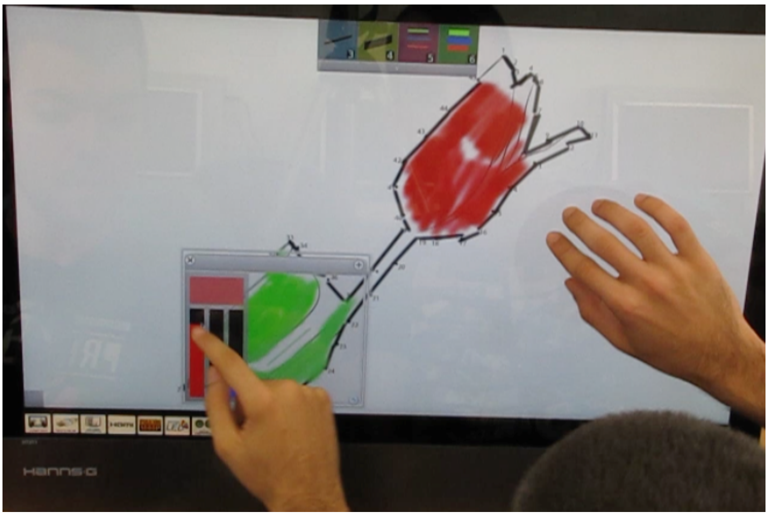


Figure 3.3: Deaf-museum: Creation stage

played in one of two modes; co-op or versus. This leaves the creation stage, in the deaf-museum application users draw using connect-the-dots guidelines. We believe, given the target demographic of pupils and students, that this stage can also be removed from the application itself and left to the teachers. The construction stage would then occur at the teacher's discretion. The benefits of this approach is reduced complexity and increased usability and short-term scalable design (Giannakos et al., 2014). This creates a new, arguably pervasive, three-stage sequence tightly coupled with the three epistemological traditions:

1. **Behaviourism** (exploration stage) - Exploring the museum

2. **Cognitivism** (application stage) - Interacting with one of the clients collaboratively

3. **Constructivism** (construction stage) - Homework/classwork left to the responsible teacher

Each stage repeats and reinforces the same topic or subject, the intention is that users visit an exhibit (1), interact with the client collaboratively (2), repeat from step 1 until the loop is over. At this point the remaining step is left to the teacher. Though not at the heart of this thesis, this repetition is interestingly similar to the spiral curriculum (Harden, 1999). This pedagogical technique requires increasing levels of difficulty, which could be incorporated into the application by increasing difficulty of questions depending on

3.1.3 Paper prototyping

As the first set of features was written down after a thorough review of related work and brainstorming, we had enough to start on the first iterations of designing the User Interface (UI). Using paper prototypes is a well established usability test and recommended first step towards designing the UI (Snyder, 2003). There are three roles in paper prototyping; facilitator, computer and subject. We're taking a slightly higher-tech approach than the one outlined by Snyder by employing a wire-framing tool: Balsamiq (Balsamiq Mockups [Computer software], 2016). We're able to integrate clicks to screen changes using this, and the role of computer as described by Snyder is replaced with the software itself.

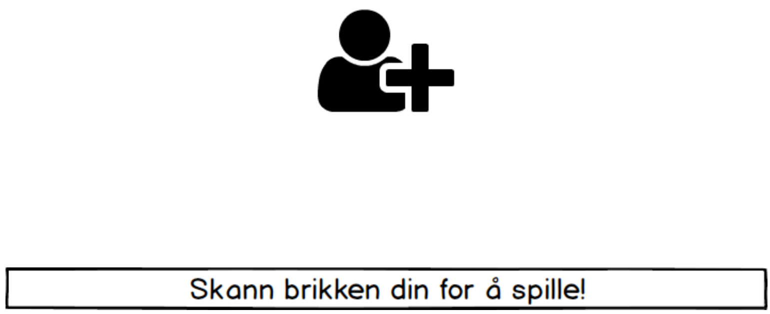


Figure 3.4: Wire-frame prototype: The log-in screen

For the complete set of images from the wire-frame prototype see Appendix 7.4.

During this phase several key design decisions were made such as the placement of the avatar, XP bar, interaction sequence and how co-op and versus mode would differ. These decisions would also inform later implementation decisions, such as uniqueness in live avatars. One major change from the original application is the multi-touch chord for answering questions during the quiz. In the original application answering left or right was done using one button, but with different gestures. After some early testing of this approach a decision was made during this prototyping to split it into two buttons, left and right. This is based on of the recommendations by Brown (1998) by adhering to user expectations. This is especially valid as obstacles in such an application is more likely to lead to session-ending frustration.

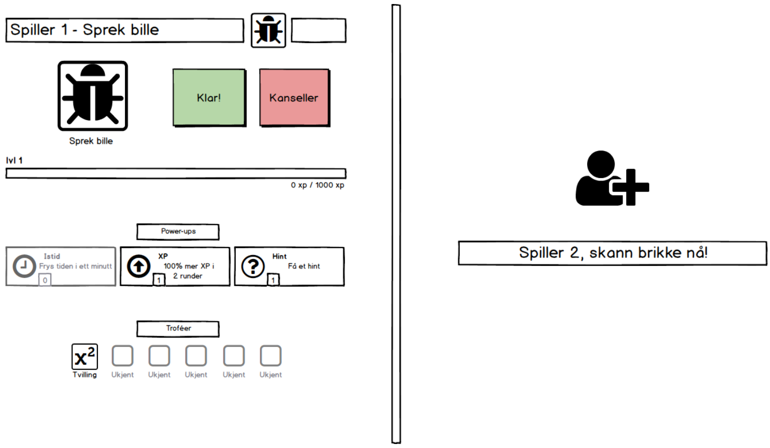


Figure 3.5: Wire-frame prototype: The player screen

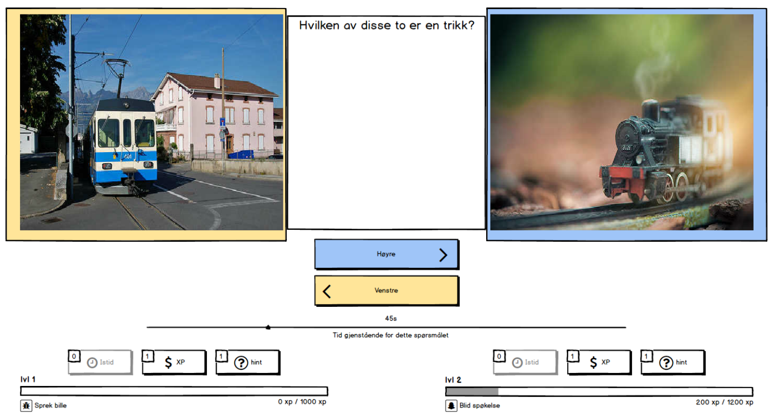


Figure 3.6: Wire-frame prototype: The co-op screen

3.1.4 Architecture and Equipment

With our prototype finalized we can discuss architecture, outline the equipment required and their interactions.

3.1.4.1 Architecture

Given the requirements of continuously updated user profiles, user authentication, and administrative tools, there is a clear-cut choice of a client-server architecture. The server will be able to store all player data and all question data. Fig. 3.7 illustrates the expected interactions.

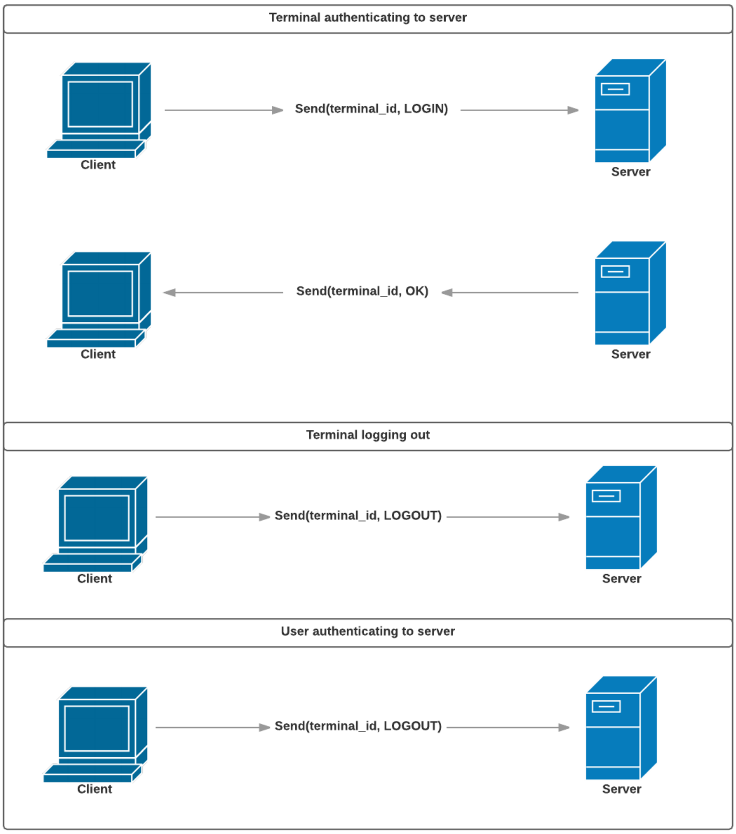


Figure 3.7: Client authentication

In this chapter we formally summarized implementation considerations and recommendations gathered in chapter 2 into requirements, divided into MUST, SHOULD and MAY categories. We then rapidly prototyped a UI and interaction design based on the original application with our new requirements in mind. We also charted out our equipment, the expected interactions between them and made a high-level architecture choice.

3.1.4.2 Equipment

The prototype equipment includes:

- Micro controller
- Proximity Coupling Device (PCD)
- Proximity Integrated Circuit Card (PICC)
- Server

- Large MT display and PC

Users will authenticate using their ascribed PICC to tap the PCD. This will then send signals to the micro-controller that processes it into a unique identifying hexadecimal string referred to as a UID. This UID shall be authenticated with the server, the specified user profile is then retrieved from the server and displayed on the large MT display.

3.2 Implementation

In this section we will discuss the implementation of our pervasive GAID application, with a working name of "peMuse". peMuse aims to engage and educate museum visitors, triggering a desire to explore the museum, collaborate or compete, and provide a low friction platform encouraging collaborative informal learning. We will first discuss the technology choices made at each step and then delve deeper into the final solution. Note: In the code-base, clients and users are termed "terminals" and "players" respectively.

3.2.1 Technology Stack

In this subsection we present the design choices and settle on a technology that is able to support those choices.

3.2.1.1 Micro-controller, PCD and PICC: Arduino UNO, RC522 RFID

Micro-controller, PCD and PICC are generic terms, the actual technologies we will be using are; Arduino Uno for the micro-controller, RFID key-fobs for the PICCs and an RC522 RFID reader for the PCD. The Arduino micro-controller will be using a modified MFRC522 library which is public domain software (Balboa, 2016). This equipment set is an inexpensive configuration using low-cost off-the-shelf components.

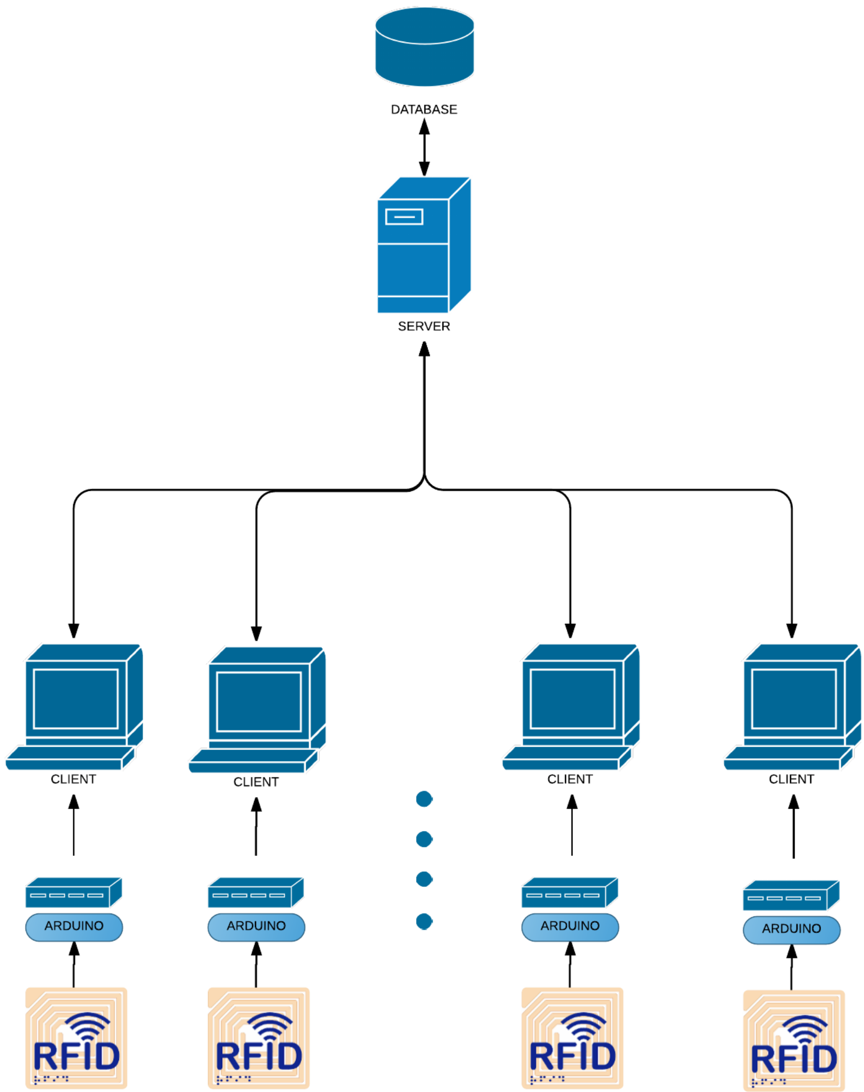


Figure 3.8: Equipment network

Figure 3.9: RFID keyfob (below) and the RC522 RFID reader

We use one arduino and one PCD for each client. Fig. 3.8 network illustrates the setup. We also have Proximity Integrated Circuit Cards (PICCs) in keyfobs with Mifare 1KB RFID embedded. These are to be given to each user and used for authentication by checking the Unique ID (UID) of the PICC against the server, see fig. 3.9.

3.2.1.2 MT client: Kivy

The client (referred to as the "terminal" in the backend) is the frontend of this system, this is the client that users interact with and employs the MT interface. The original application was written in Kivy [Computer Software] (2016), an open-source multi-platform python library. The reasons for picking this tool originally was due to the low effort involved in getting a 10-point MT prototype up and running. We pick the same tool — updated to version 1.9.1 — for similar reasons. We decided to give the prototype a working name "pe-Muse", a portmanteau of "pervasive museum", this emblazons the start-screen as shown in fig. 3.10.

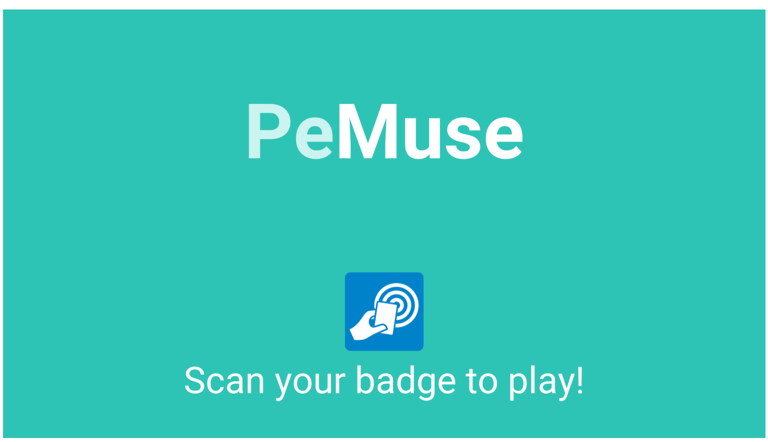


Figure 3.10: peMuse: start-screen

The Kivy implementation requires separation of concerns, presentation and logic are separated into .kv and .py files. The KV language - or sometimes kivy language - allows for widget tree creation in a declarative manner, we employ this analogously to how CSS and HTML are used in web-applications. This allows for the python code to focus on mostly logic. Given the UI from chapter 3 which transitions from screen to screen as it transitions to different contexts within the application, we decided to use kivy's screen-manager. Fig. 3.11 illustrates the inheritance from the built-in manager.

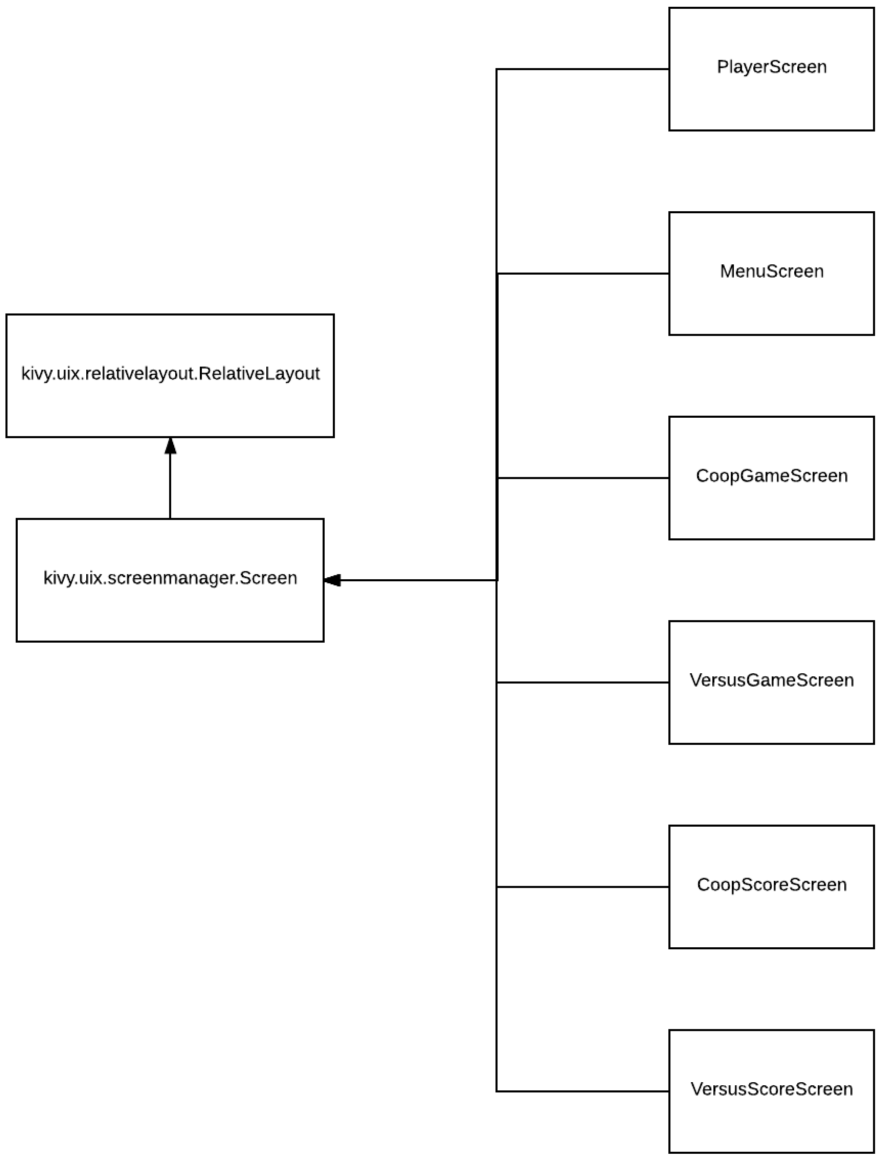


Figure 3.11: Screen-manager

Action verb	Description	API
GET	Gets all terminals	terminals/
GET	Gets terminal with private key “PK”	terminals/PK
GET	Gets player with private key “PK”	players/PK
POST	Assign new player to badge with private key “PK”	badge/PK
PUT	Set xp “XP” for player “PK”	players/PK/set_xp/XP
PUT	Set powerup quantity “Q” of certain power-up “pow_PK” for player “PK”	players/PK/set_powerup_quantity/pow_PK/Q

Table 3.1: API

```
self.sm = ScreenManager()
self.sm.add_widget(PlayerScreen(self.sm, name="player_screen"))
self.sm.add_widget(MenuScreen(self.sm, name="menu_screen"))
```

Listing 3.1: Screen-manager code snippet

The code snippet 3.1 illustrates how the screen objects are added to the screen-manager. This allows for transitions to the different screens triggered by certain events (e.g. answering a question, pressing the back button).

3.2.1.3 Server and database: Django REST framework & SQLite3

The server will be used to retrieve and set information to the player, terminal and badge model. This includes player xp, level, powerups and trophies. We chose Django REST Framework (v3.4.7) [Computer Software] (2016) (DRF) as this was familiar technology and fulfilled our requirements. The following is a subset of the API that is used in the current version of the peMuse front-end client. Note that api must have URL prefix (e.g.: http://www.example.org/) pointing at the backend server.

SQLite3 is shorthand for SQLite version 3, a SQL database engine that we employ in our DRF solution. SQLite is more than capable of handling the data load we will face in the field study, and can be kept in any eventual future large-scale field study as “Generally … any site that gets fewer than 100K hits/day should work fine with SQLite” (*Appropriate Uses For SQLite*, 2016).

3.2.2 The Solution

In this section we will take a closer look at our solution and explain relevant parts in fine detail.

3.2.2.1 Start-screen

All idle clients display the initial start-screen (see fig. 3.10), this first screen represents the second opportunity to pique a user’s interest, as the first was when they received their

PICC. This one of our external triggers (Eyal, 2014). The bright colorful display and textual prompt is our sensory stimuli, attracting and hopefully convincing users to scan their "badges" (PICC). The start-screen transitions to the right when a player has authenticated using their PICC, this reveals the players auto-generated name, level, XP, avatar and set of power-ups and their quantity.

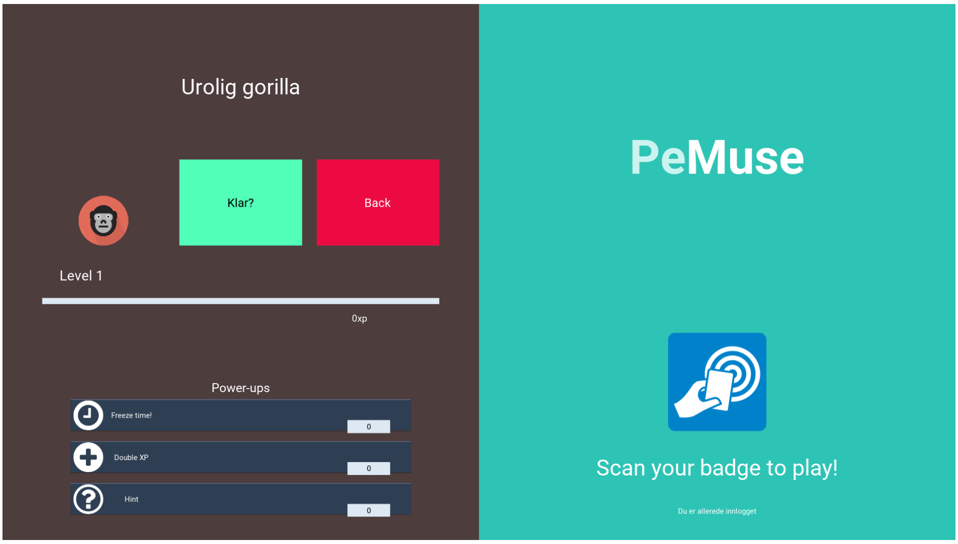


Figure 3.12: Start screen with one player (left) authenticated

We want users to feel invested and connected to their progress, and therefore link their profile with an avatar (Zaharias et al., 2013) and auto-generate animals with adjectives (e.g. hungry-hippo). This is especially useful for displaying the best players on a high-score board in a way that users themselves remember who they are on the board, but cannot identify other users. That last point is very relevant in school settings as partial anonymity can prevent abuse and dis-encouragement. Figure 3.12, 3.13 displays how players are greeted with their own avatar and profile upon authenticating with their PICC.

There are many deviations from the first brainstorming session using Octalysis, but the lessened clutter also simplifies the user interface (Giannakos et al., 2014) and makes the interaction loop easier to understand for the end-users. The top-bar from the Balsamiq prototype for instance was redundant (information overload) and removed, similar redundancy removals were made at various points from Balsamiq prototype to final solution. One exception is trophies that is removed but was not redundant, it was simply too difficult to work in given the vastly smaller scope of the experiments compared to an actual use-case with several more clients and more open-ended exploratory uses (i.e. players not necessarily playing with each other, 6+ clients, more players).

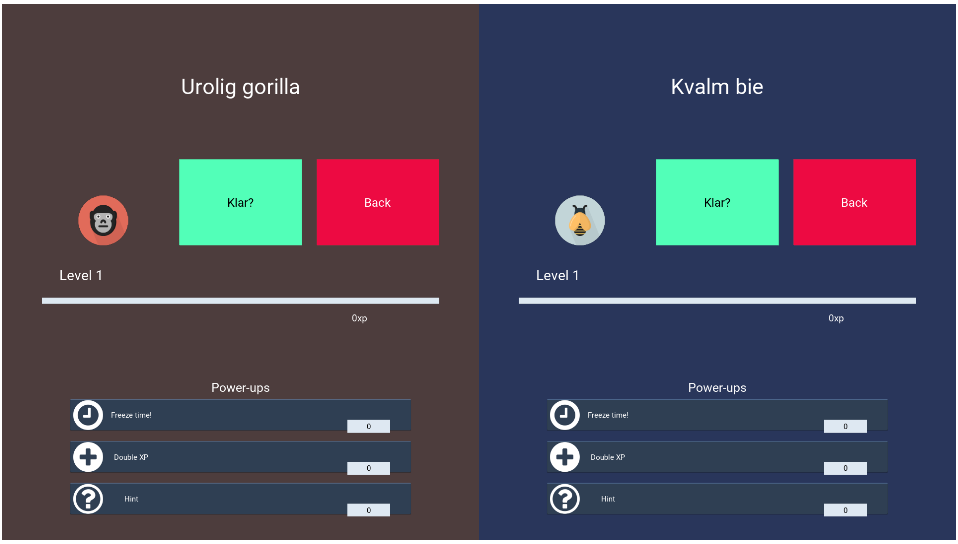


Figure 3.13: Start screen with both players

The ready and back buttons use expected colors of the common traffic light stereotype (Brown, 1998). Gamification features occupy the majority of screen space, buttons showing power-ups that popover and describe the power-up when pressed. The XP bar and level are shown in the middle of the screen. When first playing players do not know what to expect, it is therefore important to establish the game-features in order to captivate attention and pique interest in users of what's to come.

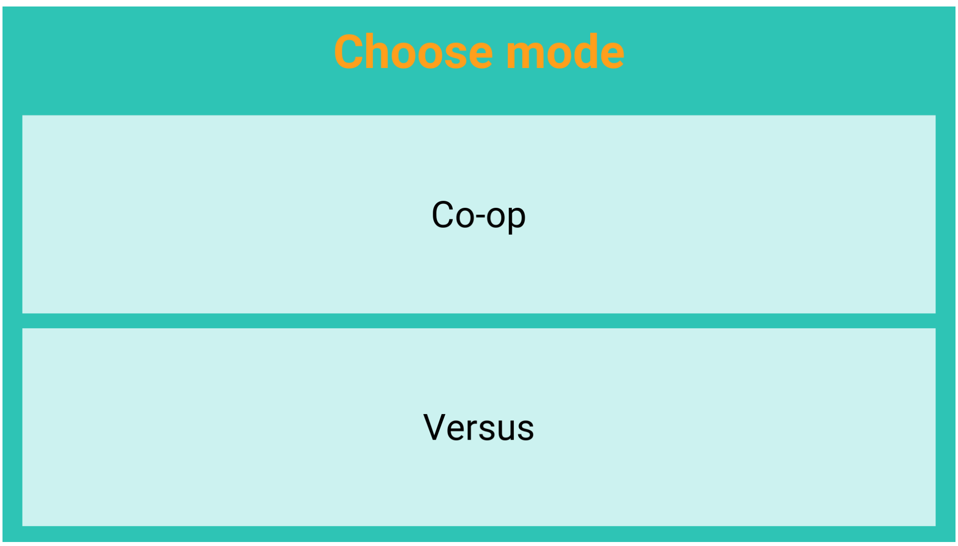


Figure 3.14: Main menu after both players are ready

When players are ready they enter a simple menu-screen (fig. 3.14) with two large buttons, coop and versus.

3.2.2.2 Co-op screen

The player's level and XP bars are displayed at all times, as well as power-ups. Power-ups that have quantity > 0 skeuomorphically appear to be click-able (shading). Main interaction point will be the central large buttons, "<< Left" and "Right >>", however early testing made clear that users press the images sometimes, for higher usability these were also made interactive for co-op mode and trigger the *on_press events in Kivy*. Figure 3.15 illustrates this screen. The light-blue line in the middle is a countdown timer, counting down from 30 seconds. One of the power-ups can stop this timer for 30 seconds. By including a timer we trigger users' loss aversion, and avoid the pitfall of multilevel games that are abandoned (Horn et al., 2012).

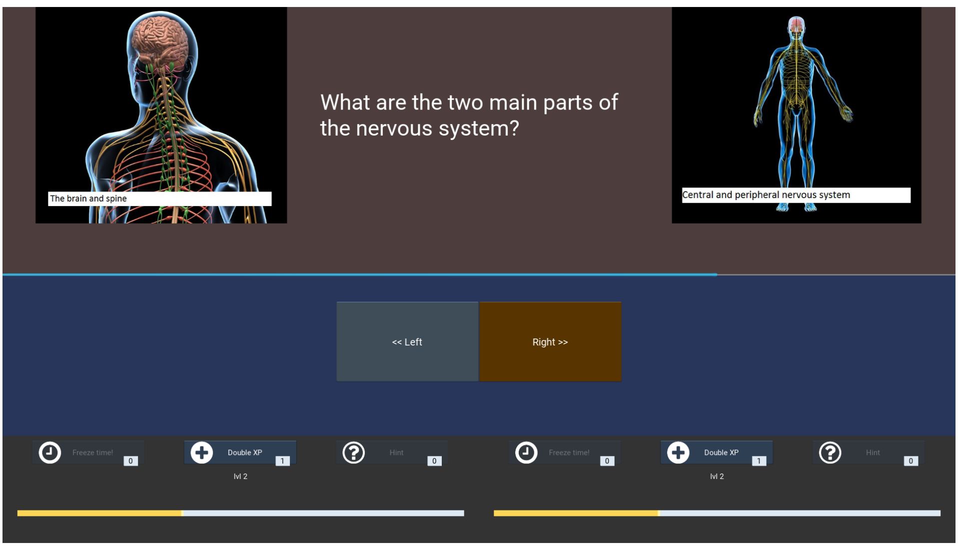


Figure 3.15: Coop screen

3.2.2.3 Versus screen

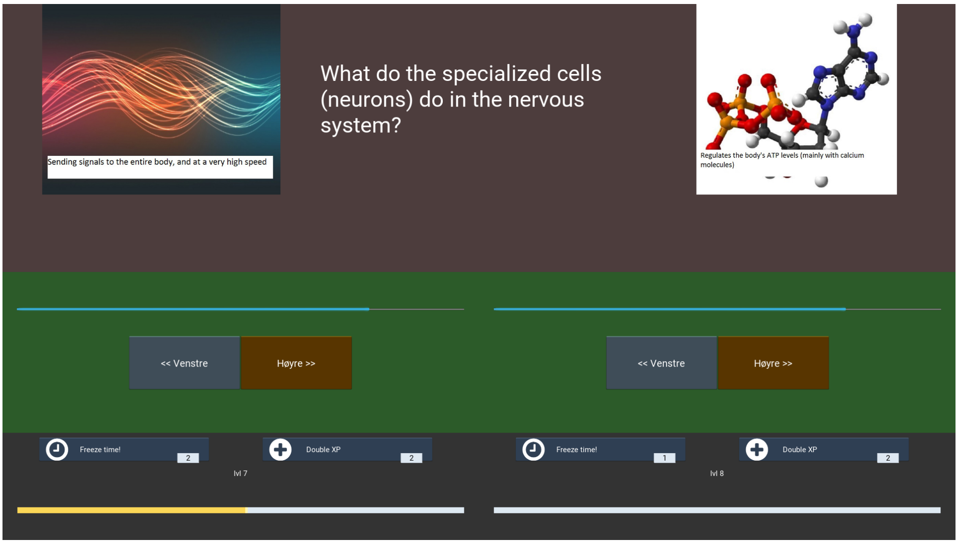


Figure 3.16: Versus screen

In versus mode the aforementioned redundancy for answering questions is disabled as the technology doesn't exist yet to distinguish different players' touch without sacrificing significant usability.

Here there are multiple sets of buttons (two pairs of Left and Right buttons), and multiple timers, otherwise this is similar to coop. There is also a slight change of background color, and we've disabled hint as it would benefit both players. Figure 3.16 illustrates this screen.

3.2.2.4 Score screen

The score screen reserves a panel for each player, coloring it green if the answer was correct and red if it was incorrect. The XP bar is animated whilst being incremented by 200 points if the answer is correct, and upon leveling up a new random power-up is immediately displayed below. This occurs after each question, and consistently displays each players score for both a competitive and a collaborative factor. Figure 3.17 illustrates this screen.

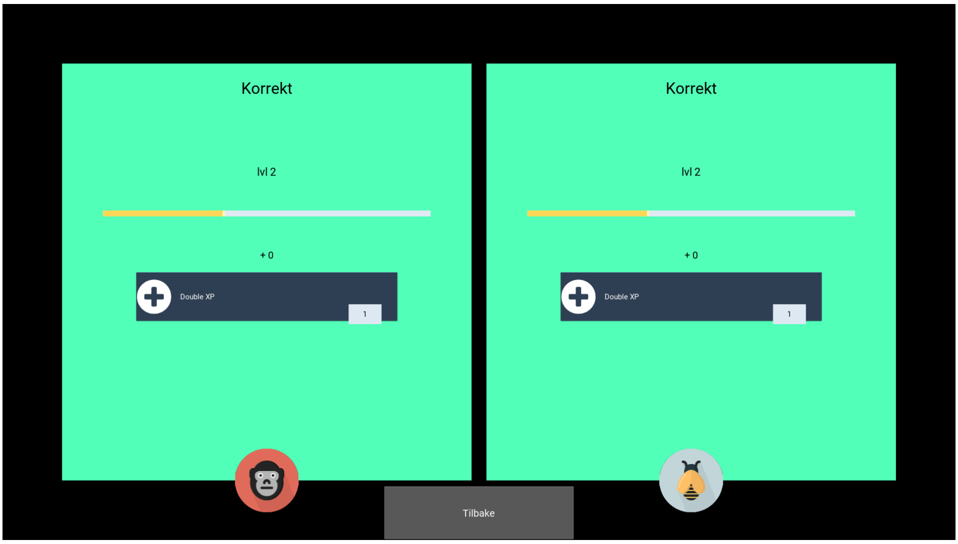


Figure 3.17: Score screen

3.2.2.5 Game rules

The game rules are as follows:

- Each question has one correct answer; left or right.
- Each question has a default timer of 30 seconds.
- Answering a question incorrectly or failing to answer in time rewards 0 XP.
- Answering correctly rewards a default of 200 XP.
- At a certain XP players level up (see Figure - XP to level function)
- Upon leveling up a random power-up is rewarded
- Power-ups are:
  - Freeze time: add 30 seconds to timer
  - Double XP: double the number of XP points from 200 XP to 400 XP.
  - Hint (disabled in versus): receive a hint for that question (3.18)

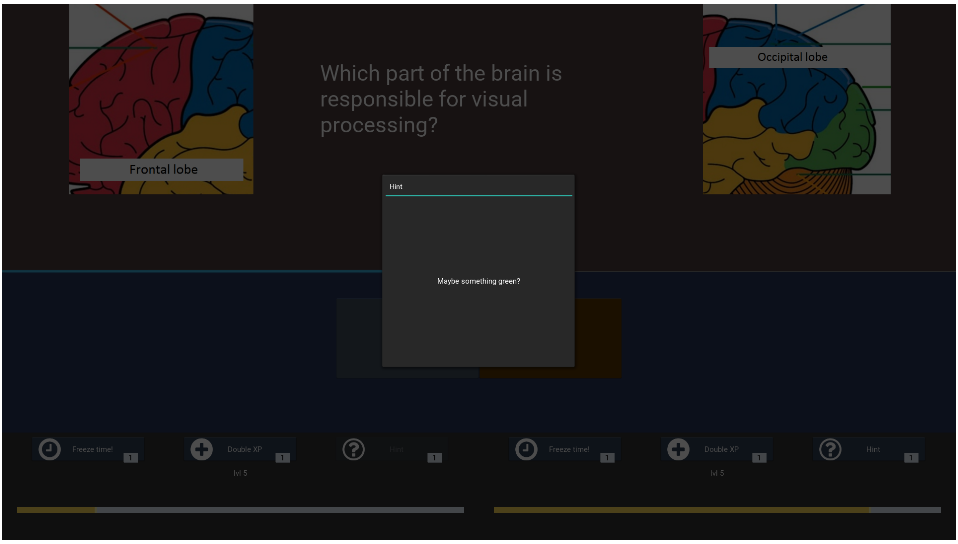


Figure 3.18: Hint being displayed

Figure 3.19 is a plot depicting our XP to level function. The Y axis represents given level and X axis the XP. There is a steadily increasing number of XP required in order to gain the next level, with the intention of prompting users to use power-ups more and making leveling up more satisfying as it becomes more rare.

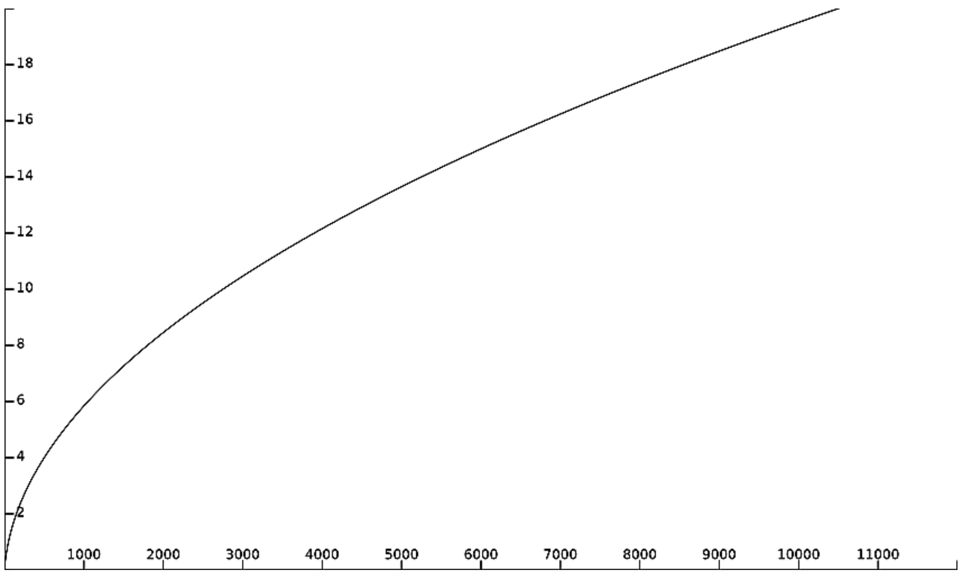


Figure 3.19: XP to level function plot, Y axis is level and X axis given XP

Chapter 4

Methodology

In this chapter we introduce our research questions, method and design of experiments.

4.1 Research questions

The goals of peMuse are to aid in both the engagement and education of museum-visitors. Our system is designed to hook users in using external triggers, establish an intrinsic motivation for them and provide a platform for rewarding informal CSCL. With this in mind, we have to develop research questions that are specific enough to determine whether peMuse achieves its goals.

We have two research questions as first introduced in chapter 1:

- RQ1: Does a pervasively gamified museum experience improve knowledge gain in comparison to a traditional museum experience?
  - Do users experience knowledge gain using the pervasively gamified system in comparison to a normal museum experience?
- RQ2: Does a pervasively gamified museum experience increase user engagement in comparison to a traditional museum experience?
  - Will users be more motivated to see and learn more, and will they be more satisfied with the gamified museum?

4.2 Research method

As mentioned in chapter 1.3, one of our research strategies was to create peMuse - our IT artifact. With this artifact as our base we set up an experiment that tests our hypothesis: peMuse will improve user engagement and knowledge gain. What kind of experiment is the best fit for testing this hypothesis?

When testing the original application (see chapter 3.1.2), the researchers employed a single-group time series design (Ross & Morrison, 1996). This is known as a quasi-experimental design and is typically employed when random assignments of subjects in control and experiment groups is infeasible (Gribbons & Herman, 1997).

Upon consulting advisers a decision was made to re-use the same repeated measurement design. Note that this does not lend itself particularly well for comparing the original application with peMuse, but for answering our research questions (i.e. knowledge gain, engagement increase). This type of quasi-experiment is convenient and often times the only viable experiment type given resources. It does however suffer additional validity threats; "The absence of randomly composed, separate experimental and control groups makes it impossible to attribute changes in the dependent measure directly to the effects of the experimental treatment" (Ross & Morrison, 1996).

When employing this experiment form care must be taken reduce validity threats, we list relevant validity threats and mitigations in table 4.1 (Oates, 2005). We omit validity threats that cannot affect the experiment.

Table 4.1: Validity threats and mitigations

Internal validity threat	Mitigation
History: interference events unnoticed between pre- and post-test observations	Video recording instruments to be used for monitoring participants.
Reactivity and experimenter effects: subjects altering behaviour due to being observed	Significant number of participants are acquaintances and friends, blind experiments would control for this. Care is taken to give subjects as little information about the experiment, especially with regards to the domain (neurology) as knowledge gain is an important metric to test for.
External validity threat	
Over-reliance on special types of participants	All participants are students, posing a threat to validity which cannot be mitigated. Another threat is the volunteer nature of picking participants which might pose a threat as well (Cudrin, 1969).
Too few participants	High enough number of participants: 36

4.3 Research design

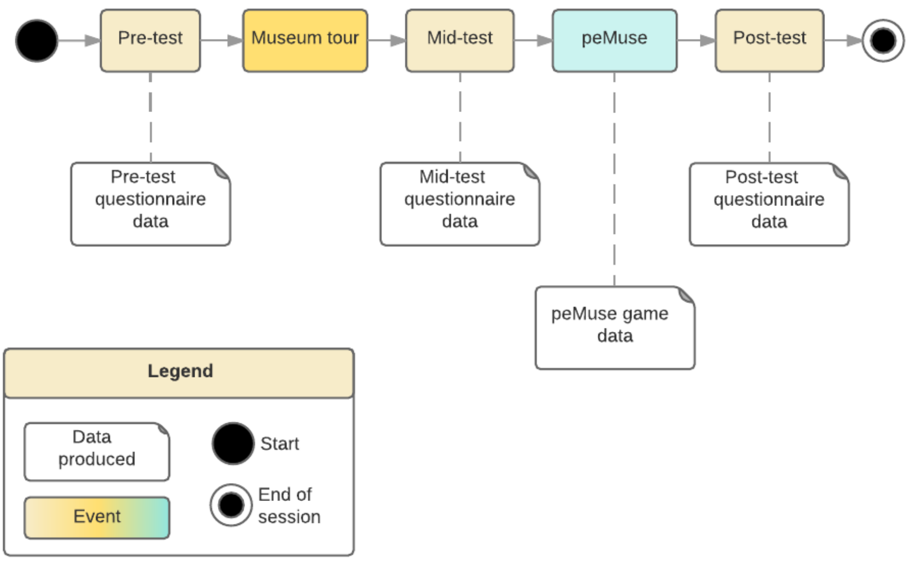


Figure 4.1: Experiment state diagram

Figure 4.1 illustrates the single-group time series design and when our pre-, mid- and post-tests occur. This design involves "...repeated measurement of a group with the experimental treatment induced between two of the measures" (Ross & Morrison, 1996). Our experiment will feature pairs of users which we refer to as dyads. The experiment will take place in a neutral room at the university campus with posters designed to emulate a museum. The users will be briefed and asked to conduct a pre-test. This pre-test will help establish a knowledge baseline (how much do they know about the subject at hand). After this point they will be equipped with eye-tracking goggles and asked to browse the room as if it were a museum (museum tour). After this phase there is another observation before the treatment which we call the mid-test. After the users have finished the mid-test they are given their badges and given the treatment; using the peMuse system with their PICC. Finally, they are to be observed a final time with the post-test.

4.3.1 Questionnaires

The pre-, mid- and post-test questionnaires are to be found in appendices 7.1, 7.2 and 7.3.

The pre-test establishes gender, age and prior experience with technology, and small and large MT devices. It then establishes prior knowledge of our domain, neurology, via a 10 question survey. Here participants are told to not answer if they aren't sure in order to keep our estimates as accurate as possible.

The mid-test questionnaire consists of 10 questions and has the same theme of questions. This test occurs after the participant pair has browsed the museum, but before treatment with our application.

The post-test questionnaire also has 10 questions, and includes a Likert (1932) based survey at the end with 11 sections (table 4.2).

Table 4.2: Factors used in survey

Factors	Operational Definition	# of items / questions
Satisfaction	Degree of satisfaction with enhanced activity.	3
Intention	Degree of positive future intention to repeat the enhanced activity in the future.	4
Easiness	Degree of belief that the enhanced activity is easy.	4
Enjoyment	Degree of enjoyment with the enhanced activity.	4
Usefulness	Degree of usefulness of the enhanced activity.	4
Control	Degree of belief that user is in control of the enhanced activity.	2
Characterization (activity)	How user characterizes the enhanced activity.	3
Characterization (performance)	How user characterizes performance with the enhanced activity.	6

4.3.2 Domain choice

The domain chosen was neurology, as such posters, questionnaires and game content was fit to this. One of the advisors had a substantial domain expertise in this field and provided input on the design of the posters and questions for questionnaires.

4.3.3 Data generation

As shown in fig 4.1, each questionnaire produces quantitative data. The peMuse application also sends player scores, power-ups, level, xp, questions answered, elapsed time in question answering and player played with to the server. Retrieving this data from the database in the form of a JSON allows another set of data.

Our data generation methods will consist of observations, questionnaires from pre- mid- and post-tests, as well as the peMuse application game data.

4.3.4 Dependent variables

We normalize test scores to be between 0 and 1. Our two dependent variables are scores from mid- and post-test.

4.3.5 Process variables

We regard the individual experience points (XP) as our game performance index.

4.3.6 Experiment setup

he field study took place on the 21st and 24th of October as the eye-tracking equipment was available those days. The participants were recruited via social media, direct messages, and recruitment via advisors. They numbered 36 participants in 18 dyads. There were 13 females and 23 males with an average age of 24.4 years with a standard deviation of 5 years. Given the informal CSCL narrative, we explicitly encourage participants to collaborate 4.2.

Figure 4.2: Experiment pictures: (a) a dyad viewing a poster, and the posters in the experiment room

After the museum tour and mid-test each dyad was instructed to use their PICC to start the peMuse gamified quiz, once collaborative (coop) and once competitive (versus). Figure 4.3 shows one of the clients (a) in idle mode waiting for a PICC to scan on the PCD (red). In (b) we see one of the dyads playing competitive mode.

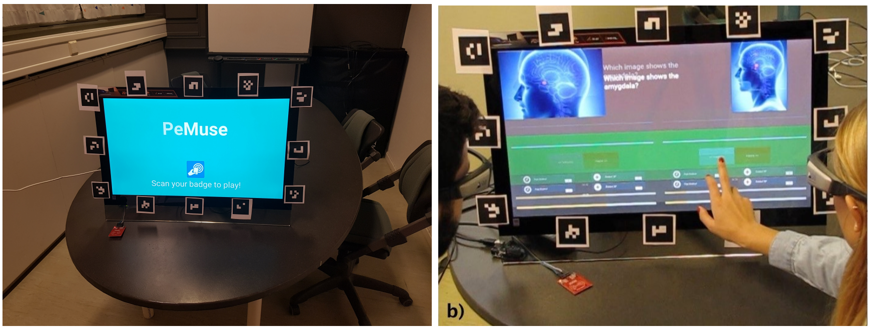


Figure 4.3: Experiment pictures: peMuse — one of two clients idle, and (b) participants playing

After completing the final terminal both participants are asked to fill out a final post-test and rewarded with the equivalent of NOK 85.

Chapter 5

Results

In this chapter we will present the results from the field study. We set our significance level to the common P ¡ 0.05 (Fisher et al., 1950).

5.1 Pre-test and mid-test: Improvement in knowledge

Improvement in knowledge from pre-test to mid-test was significant. Fig. 5.1 The pre-test established previous knowledge level of the domain (neurology) that our museum-proxy was based on. The scores in the mid-test are significantly higher than the pre-test.

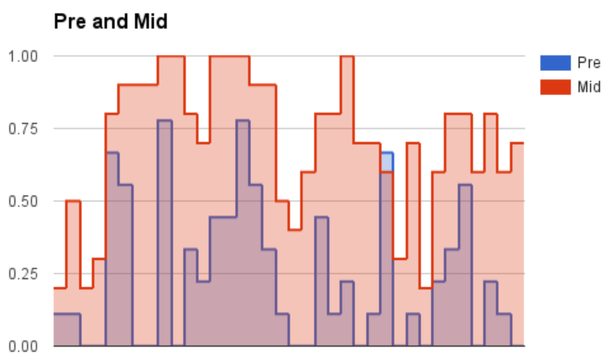


Figure 5.1: Graph illustrating gain from pre- to mid-test

Table 5.1: Participants and their score (dyads are same greytone)

Participant	Pre-test score	Mid-test score	Pre-Mid Difference	Post-test score	Mid-Post Difference
E	0.11	0.20	0.09	0.33	0.13
I	0.11	0.50	0.39	0.89	0.39
Tb	0.00	0.20	0.20	0.22	0.02
Tr	0.00	0.30	0.30	0.33	0.03
sheep	0.67	0.80	0.13	0.89	0.09
spider	0.56	0.90	0.34	0.78	-0.12
bear	0.00	0.90	0.90	0.89	-0.01
turtle	0.00	0.90	0.90	0.89	-0.01
giraffe	0.78	1.00	0.22	0.78	-0.22
koala	0.00	1.00	1.00	0.89	-0.11
leopard	0.33	0.80	0.47	1.00	0.20
bee	0.22	0.70	0.48	0.67	-0.03
tiger	0.44	1.00	0.56	0.89	-0.11
bee	0.44	1.00	0.56	0.89	-0.11
chicken	0.78	1.00	0.22	0.78	-0.22
leopard	0.56	0.90	0.34	0.89	-0.01
cat	0.33	0.90	0.57	0.89	-0.01
dolphin	0.11	0.50	0.39	0.33	-0.17
tiger	0.00	0.40	0.40	0.11	-0.29
bat	0.00	0.60	0.60	0.44	-0.16
leopard	0.44	0.80	0.36	0.89	0.09
tiger	0.11	0.80	0.69	0.78	-0.02
chicken	0.22	1.00	0.78	1.00	0.00
duck	0.00	0.70	0.70	0.56	-0.14
zebra	0.11	0.70	0.59	0.78	0.08
squirrel	0.67	0.60	-0.07	0.78	0.18
fox	0.00	0.30	0.30	1.00	0.70
penguin	0.11	0.70	0.59	0.89	0.19
monkey	0.00	0.20	0.20	0.56	0.36
shark	0.22	0.60	0.38	0.67	0.07
bee	0.33	0.80	0.47	0.89	0.09
mus	0.56	0.80	0.24	0.89	0.09
gorilla	0.00	0.60	0.60	0.56	-0.04
crocodile	0.22	0.80	0.58	0.78	-0.02
panda	0.11	0.60	0.49	0.44	-0.16
cat	0.00	0.70	0.70	0.78	0.08
AVG	0.24	0.70	0.46	0.72	0.02

We measure significance of our observations using paired t-tests. We have dependant samples — we collect multiple samples on each participant that are closely related. This outputs paired observations (e.g. before and after) for each observation, making paired t-test a good choice.

Table 5.2: Testing knowledge gain from pre-test to mid-test (before treatment). Standard Deviation (SD), $p < 0.05$

	Mean (SD)		T
Learning performance	**Pre**	**Mid**	
	0.24 (0.251)	0.70 (0.245)	11.63*

S.D. Standard Deviation; $*p < 0.05$

We simply calculate using paired t-test for our difference mean $\overline{d} = 0.462345679$ and standard deviation $S_d = 0.238458864$.

$$SE(\overline{d}) = \frac{S_d}{\sqrt{n}} = \frac{0.238458864}{\sqrt{36}} = 0.039743144$$

$$T = \frac{\overline{d}}{SE(\overline{d})} = \frac{0.462345679}{0.039743144} \approx 11.63$$

This T-value gives us P value < 0.00001, which is statistically significant for our P value (0.05).

5.2 Mid- to Post-test: No improvement in knowledge

There was no statistically significant improvement from mid-test to the post-test. Note that the mid-test occurs before treatment (i.e. peMuse application) but after the museum tour. See table 5.1 and table 5.3.

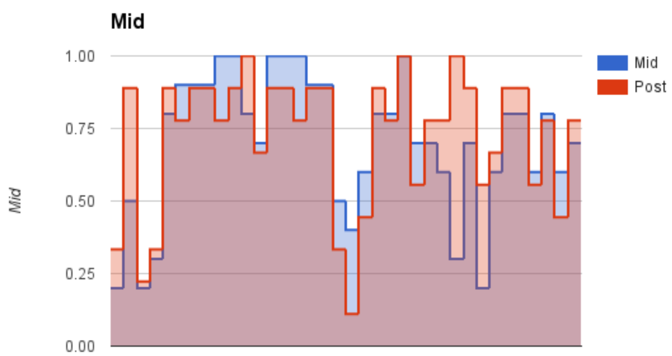


Figure 5.2: Graph illustrating gain from mid- to post-test

Table 5.3: Testing knowledge gain from pre-test to mid-test (after treatment)

	Mean (S.D.)		T
	Pre	**Mid**	
Learning performance	0.70 (0.245)	0.72 (0.235)	0.70

S.D. Standard Deviation; $*p < 0.05$

Mean $\overline{d} \approx 0.022$ and standard deviation $S_d = 0.1897738374$

$$SE(\overline{d}) = \frac{S_d}{\sqrt{n}} = \frac{0.1897738374}{\sqrt{36}} = 0.03162897289$$

$$T = \frac{\overline{d}}{SE(\overline{d})} = \frac{0.02222222222}{0.03162897289} \approx 0.70$$

This T-value gives us P value ≈ .24, which is larger than our P value (0.05), resulting in no significant difference for this paired T-test. A Shapiro test reveals no normality in differences, and Wilcoxon also shows no significant difference (Appendix 7.6. There is however a large outlier from the tiger participant in the tiger-bat dyad in the post-test, managing only 11% correct answers, but 40% before in the mid-test. Figure 5.3 illustrates the relationship between pre- mid- and post-test clearer (note: Posttest1 is the mid-test). Omitting this dyad does not change the conclusion with regards to statistical significance.

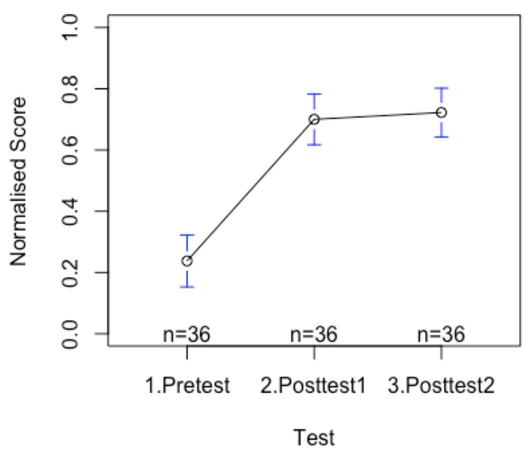


Figure 5.3: Test scores, points show the mean values among all participants and the blue bars show the 95% confidence intervals

5.3 Survey results

Results gathered indicate users were very satisfied with the experience and enjoyed the experience. Several users optionally writing in that the experience was fun and both usefulness and satisfaction scores being high. Participants also display enjoyment in the video, with many laughing at many points. The data gathered (Table 5.4) and calculations (again in paired T-tests) indicate that users score higher in all fields with a statistically significant margin (fig. 5.4).

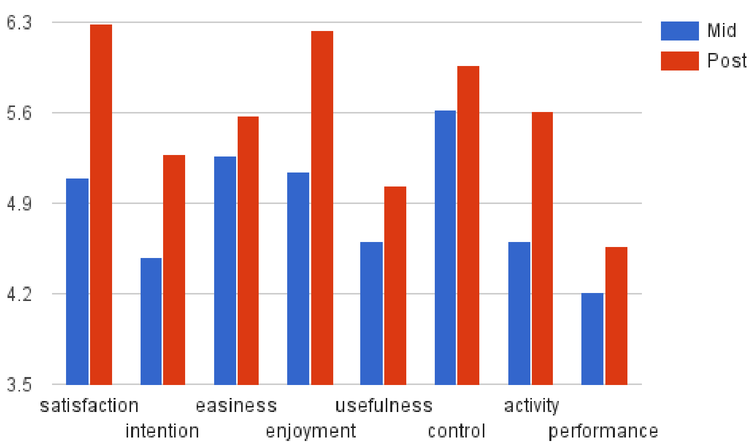


Figure 5.4: Graph illustrating mean difference using our Likert-scale from mid- to post-test. Categories are all statistically significant (Appendix 7.5.

Table 5.4: Data from the Likert survey

	Mean (S.D.) Pre	Mean (S.D.) Mid	T	P
Satisfaction	5.10 (1.35)	6.29 (0.77)	4.89569317*	0.000011
Intention	4.49 (1.40)	5.28 (1.37)	3.71897681*	0.000349
Easiness	5.27 (1.00)	5.58 (1.06)	2.01794831*	0.025659
Enjoyment	5.15 (1.50)	6.24 (0.78)	4.12802746*	0.000108
Usefulness	4.60 (1.30)	5.03 (1.30)	2.76389394*	0.004525
Control	5.63 (1.20)	5.97 (1.01)	1.71897736*	0.047234
Activity	4.61 (1.23)	5.61 (0.88)	4.81070235*	0.000014
Performance	4.21 (0.56)	4.58 (0.51)	3.5824529*	0.000512

S.D. Standard Deviation; $*p < 0.05$

5.4 Hedonic/Utilitarian attitude measure

We also employed the two-dimensional Hedonic/Utilitiarian attitude measuring technique (Voss et al., 2003). The questions can be found in the Appendix 7.3.

Table 5.5: Measure of the Hedonic and Utilitarian dimensions

Utilitarian	**Mean (S.D.)**	**Mean Utilitarian (S.D.)**	T	P
MT_effective	6.06 (0.95)	5.72 (0.91)	2.428690648*	0.010217
MT_helpful	5.72 (1.19)			
MT_functional	6.00 (0.93)			
MT_practical	6.03 (0.97)			
MT_necessary	4.78 (1.76)			
Hedonic		**Mean Hedonic (S.D.)**		
MT_fun	6.47 (0.88)	6.03 (0.87)		
MT_exciting	6.14 (1.05)			
MT_delightful	5.75 (1.02)			
MT_enjoyable	6.22 (0.99)			
MT_thrilling	5.56 (1.27)			

S.D. Standard Deviation; $*p < 0.05$

We must mention pre-test question 3 and post-test question 4 were incorrect, this oversight reduces pre- and post-test to 9 questions in total. These questions are totally excluded from our results.

5.5 Qualitative input

In our post-test we featured two suggestion boxes for free input from our participants. The prompts were:

Do you have any suggestions in order to improve the activity?

and

Here you can write whatever you want for the activity:

Table 5.6: Suggestions (SG) and Anything Goes (AG) gathered from the surveys

SG1	show the right answer after the persons answered. It was easy to forgetwhat was left and what was right
SG2	get score based on time in VS mode. More aggressive tools in VS mode (e.g. 1/2 time of opponent)
SG3	make it harder to see what the other person is answering. More information about the game mechanics before we start.
SG4	more flashy leaderboard. More alternatives
SG5	fix some small bugs
SG6	english language in tests
SG7	Get an english version of poster and games
SG8	explain the rules better
SG9	give points for answering quickly
AG1	it was fun
AG2	it was fun
AG3	most of the enjoyment was due to the social aspect
AG4	it was very fun
AG5	Great
AG6	highscore list
AG7	LIT!
AG8	Like the idea of choosing a subject that can have application in the real life, plus able to learn sth more about how our body functions
AG9	bra jobba (translated: good job)

This is an exhaustive list and features all input from users that wanted to leave input.

5.6 Game results

Our application also produces data we can present here, each session users answer a series of questions granting them XP, power-ups, and levels. We use XP as a performance score.

We observe three significant correlations

- Scores in the first and mid-test are correlated $(r(34) = 0.69, p < 0.0001)$. This means that participants that scored high in the mid-test also score high in the post-test (Figure 5.5 (a))
- The game score is correlated to the score in the mid-test $(r(34) = 0.42, p = .01)$. This means that the participants that score high in the mid-test also score high in the game (Figure 5.5 (b)).
- The game score is correlated to the score in the post-test $(r(34) = 0.34, p = .04)$. This means that the participants who perform well in the game also score high in the post-test.

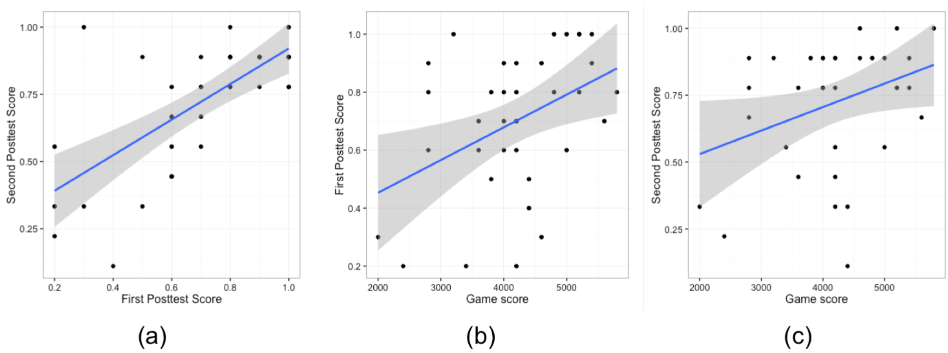


Figure 5.5: Scatter plot, (a) mid-test and post-test score, (b) game score and mid-test score, and (c) game score and post-test score. The blue line indicates the linear model for the variable on the Y-axis given the variable on the X-axis. The grey area shows 95% confidence interval.

Chapter 6

Discussion

In the last chapter we presented the results from our experiments, in this chapter we will discuss what our findings reveal. We will reiterate previous research and compare it to our findings, and in which ways our findings either agrees with, extends, refines or conflicts with the previous research.

6.1 Statistically insignificant improvement to knowledge gain

One of our research questions (RQ1) was: "Does a pervasively gamified museum experience improve knowledge gain in comparison to a traditional museum experience?". Users overwhelmingly gained a higher knowledge level of the domain material after browsing our proxy-museum, as was expected. The pre-test to mid-test scores were significantly higher, unfortunately the mid-test to post-test scores did not experience a similar jump. Although there was a slight improvement, it was not statistically significant with our moderate $P < 0.05$. Why did this occur? In order to attempt an answer to this question, we must investigate our next RQ.

6.2 Improved user engagement

Our second research question (RQ2) was: "Does a pervasively gamified museum experience increase user engagement in comparison to a traditional museum experience?". For this research question our research design allowed for both quantitative and qualitative data. Using an 8-part scale (Likert, 1932) we surveyed our participants satisfaction, intention, easiness, enjoyment, usefulness, control, activity and performance (see table 4.2). From the results we can conclude that participants scored high on each point. Only for easiness and control does the P value go above 0.02. Participants were also positive in the qualitative portions of our experiment. Table 5.6 collects all the feedback and we divide them by suggestions (SG) and Anything Goes (AG). Participants were not required to answer this and for this reason we only have nine items for each field. Note that SG_a and AG_a does not necessarily source from the same participant.

As a bonus we also added a Hedonic/Utilitarian attitude measure (Voss et al., 2003) (table 5.5). Results indicate that our application leans towards hedonic tendencies, participants associated peMuse with fun more than practical (though both scores are high).

6.3 Why knowledge gain was insignificant

SG5 states that bugs were present, and upon reviewing the field-study videos it became clear that bugs were present for a significant portion of the participants. The bugs were primarily due to wrong URLs for the question-images. Each question is stored in the database and retrieved by each client on start-up. Some of the question objects had a non-existent URL, which lead to a continuous loading spinner where the question-image should have loaded. Another bug — though less prominent — was misspellings of certain questions. There is no doubt that this poses a slight validity threat, and should be a concern when coupled with the inherent validity threats of the single-group time series design (Ross & Morrison, 1996).

SG7 and SG6 comes from a group that was scheduled incorrectly, as the experiment took place over two days and the first day consisted of Norwegian students. SG8 is the final suggestion that highlights another potential validity threat, rules of the game were not satisfactorily explained. Our application should prompt first-time participants/users and explain the rules of the game. SG8 ties into our relatively lowered improvement in control (table 4.2). Cognitive affordance increase (Hartson, 2003) coupled with a first-time tutorial approach should solve these issues, and is recommended for final solutions. SG1, SG2, SG3 and SG9 suggest design-improvements for peMuse and should be considered in any future work. Perhaps most crucially SG1's suggestion — as this might single-handedly contribute to a larger increase in knowledge gain post-treatment. This resurfacing of previously explained concepts fits in well with the spiral curriculum method (Harden, 1999).

The issues highlighted by participants helps narrow our search effort for the cause of the insignificant knowledge gain. Our application has some improvement points — some suggestions are even featured in our requirements (albeit under MAY 3.1.1.3) such as high-scores (SG4, AG4). Our application should also aim to be bug-free and multiple pilot-studies should be run. However, participants still overwhelmingly favoured the application in comparison to the proxy-museum. Sentiment was generally positive, and can be seen in feedback gathered from the text-fields as well (AG1, AG2, AG4, AG5, AG7, AG9).

6.4 Research critique

Our choice of the quasi-experimental single-group time series design was justified given the resources available. The validity threats that are inherent to this design can be dealt with, but coupled with our applications bugs and above-mentioned field-study hiccups, the validity threat is a concern. This study could also have used another set of data from the game besides XP had it been captured by the application, for instance every interaction with the screen or every power-up usage. These statistics would provide more information that could inform our otherwise speculative inquiry. One final criticism centres on the venue. Our proxy-museum did receive praise (though mostly for domain choice), but the participant numbers would be increased and the Hawthorne effect (McCarney et al., 2007) decreased. The target venue is also beneficial to execute field studies in comparison to a proxy-museum, domain-expert deficiency aside. A decision was made earlier to not conduct the experiment in a museum however, we believe that a minimum of 6 clients is necessary in order to capture the above-mentioned more meaningful data. The pervasiveness of this system truly shines when scaled up.

Chapter 7

Conclusion

Our research had the goal of investigating whether a pervasively gamified museum can positively contribute with regards to knowledge gain and increased user engagement. This thesis originally started with the goal of improving the gamification features of an existing application. Upon reviewing the available literature, a decision was made to design a new application from the ground up with an innovative approach of pervasively gamifying a museum.

In order to answer these research questions we devised a two-pronged research strategy; design and creation of an IT artifact (peMuse) and subsequent experiments. Our field-study resulted in two conclusions: (1) knowledge gain is insignificantly increased, and (2) user engagement is significantly increased. Qualitative data also reinforced the positive RQ2 results. In the previous chapter we discussed and identified the potential causes for the negative RQ1 results.

Given our mostly positive results we believe future work is warranted. Our application presents solutions to problems regarding lack of oversight (especially for teachers on field trips with many pupils), self-assessment, individualized feedback and increased motivation to explore more. peMuse isn't without faults however, the previous chapter discusses our pitfalls, and we now present advice for future work regarding both the application and research design.

Future work

As mentioned in the previous chapter, the quasi-experimental design was a short-coming. Given the required resources, switching to true experiments would reduce validity threats significantly. In addition, as Seaborn & Fels (2015) suggests: "One way that the effect of gamification can be measured – especially longitudinally – is by removing gamification features from the system". In this section we will list our suggestions for future work (both application and research).

- Prototyping should have leaned more towards user-centered design, future iterations have a great opportunity given modifiability of current application.

- More justification for picking Kivy. Although easy to get started, many similar but web-based frameworks could provide similar functional benefits such as MT and a more familiar HTML-CSS-JS separation.
- In field study analysis a lot of time could have been shaved off had every interaction with screen been stored, for instance as a PUT request to the server with which button was pressed.
- Decision to drop trophies due to limited scope of the field study, particularly the number of clients and size of the proxy-museum.
- Question model should have left_image_text and right_image_text instead of hard-copying the text onto images during their creation.
- With the data gathered a high score screen could be situated several places in the museum or on idle client screens. Promotes competitive behaviour, user engagement and not feeling left out.
- Instead of abandoning the creation stage (see original application 3.1.2), or leaving it to the teacher/parent, a secondary mode within the game focusing on creation could be incorporated. Instead of solely relying on a quiz game in traditional stage. Leaving the application as a quiz might lead to unwanted simplicity and reduced engagement when scaled up to many clients.
- The pervasiveness of our system specifically allows for users to pick and choose which exhibits they want to visit, making the experience open-ended (Horn et al., 2012). This requires a larger scope in order to be taken advantage of, the following measures are recommended take advantage of the pervasiveness of peMuse:
  - Multiple clients (6+)
  - Progress bar for which exhibits' client have been completed
  - Participants more random distribution of population - no high concentration of university students as in our experiment (Oates, 2005).
  - At an actual museum.
- Motivating visitors to see more of the museum (completion-rate) is an intended side-effect of this system. Exploring this further could yield interesting answers.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, *37*(4), 445–456.

Amabile, T. M., Goldfarb, P., & Brackfleld, S. C. (1990). Social influences on creativity: Evaluation, coaction, and surveillance. *Creativity research journal*, *3*(1), 6–21.

Appropriate uses for sqlite. (2016). n.d. Retrieved December 15, 2016, from, from `https://sqlite.org/whentouse.html`

Balboa, M. (2016). *Miguelbalboa/rfid.* Retrieved November 21, 2016, from, from `https://github.com/miguelbalboa/rfid`

Balsamiq Mockups [Computer software]. (2016). Retrieved from, from `https://balsamiq.com/`

Benford, S., Magerkurth, C., & Ljungstrand, P. (2005). Bridging the physical and digital in pervasive gaming. *Communications of the ACM*, *48*(3), 54–57.

Bradner, S. (1997). Key words for use in rfcs to indicate requirement levels.

Brown, C. M. (1998). *Human-computer interface design guidelines*. Intellect Books.

Chew, C. B., Mahinderjit-Singh, M., Wei, K. C., Sheng, T. W., Husin, M. H., & Malim, N. H. A. H. (2015). Sensors-enabled smart attendance systems using nfc and rfid technologies. *International Journal of New Computer Architectures and Their Applications*, *5*(1), 19–29.

Chou, Y.-K. (2013). Octalysis: Complete gamification framework. *Yu-Kai Chou & Gamification*.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*, *125*(6), 627.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic mindtrek conference: Envisioning future media environments* (pp. 9–15).

Django REST Framework (v3.4.7) [Computer Software]. (2016). Retrieved from, from `http://www.django-rest-framework.org/`

DomíNguez, A., Saenz-De-Navarrete, J., De-Marcos, L., FernáNdez-Sanz, L., PagéS, C., & MartíNez-HerráIz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, *63*, 380–392.

Eyal, N. (2014). *Hooked: How to build habit-forming products*. Penguin Canada.

Fisher, R. A., et al. (1950). Statistical methods for research workers. biological monographs and manuals. no. v. *Statistical methods for research workers. Biological monographs and manuals. No. V.*(11th ed.).

Fogg, B. J. (2002). Persuasive technology: using computers to change what we think and do. *Ubiquity*, *2002*(December), 5.

Giannakos, M. N., Jones, D., Crompton, H., & Chrisochoides, N. (2014). Designing playful games and applications to support science centers learning activities. In *International conference on universal access in human-computer interaction* (pp. 561–570).

Google Trends. (2016). *Search term "gamification"*. Retrieved from, from `https://www.google.com/trends/`

Gribbons, B., & Herman, J. (1997). *True and quasi-experimental designs*. ERIC Clearinghouse on Assessment and Evaluation.

Harden, R. M. (1999). What is a spiral curriculum? *Medical teacher*, *21*(2), 141–143.

Hartson, R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology*, *22*(5), 315–338.

Horn, M., Atrash Leong, Z., Block, F., Diamond, J., Evans, E. M., Phillips, B., & Shen, C. (2012). Of bats and apes: an interactive tabletop game for natural history museums. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 2059–2068).

Huotari, K., & Hamari, J. (2012). Defining gamification: a service marketing perspective. In *Proceeding of the 16th international academic mindtrek conference* (pp. 17–22).

Kivy [Computer Software]. (2016). Retrieved from, from `https://kivy.org/`

Lee, J. J., & Hammer, J. (2011). Gamification in education: What, how, why bother? *Academic exchange quarterly*, *15*(2), 146.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.

Linehan, C., Kirman, B., Lawson, S., & Chan, G. (2011). Practical, appropriate, empirically-validated guidelines for designing educational games. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1979–1988).

McCarney, R., Warner, J., Iliffe, S., Van Haselen, R., Griffin, M., & Fisher, P. (2007). The hawthorne effect: a randomised, controlled trial. *BMC medical research methodology*, *7*(1), 1.

Oates, B. J. (2005). *Researching information systems and computing*. Sage.

Prince, D. R. (1990). Factors influencing museum visits: An empirical evaluation of audience selection. *Museum Management and Curatorship*, *9*(2), 149–168.

P&S Market Research. (2016). *Gamification market growing at 46.3% cagr to 2020 led by social media.* Retrieved November 24, 2016, from, from http://www.prnewswire.com/news-releases/gamification-market-growing-at-463-cagr-to-2020-led-by-social-media-568923551.html

Ross, S. M., & Morrison, G. R. (1996). Experimental research methods. *Handbook of research for educational communications and technology: A project of the association for educational communications and technology*, 1148–1170.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31.

Snyder, C. (2003). *Paper prototyping: The fast and easy way to design and refine user interfaces*. Morgan Kaufmann.

Susi, T., Johannesson, M., & Backlund, P. (2007). Serious games: An overview.

The Government of Canada. (2004). *2004 survey of visitors to museums' web space and physical space: Survey documentation and findings.* Retrieved from http://canada.pch.gc.ca/eng/1443530283817/1443530281438

Vaishnavi, V. K., & Kuechler, W. (2015). *Design science research methods and patterns: innovating information and communication technology*. Crc Press.

Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the hedonic and utilitarian dimensions of consumer attitude. *Journal of marketing research*, *40*(3), 310–320.

Watson, D., Hancock, M., Mandryk, R. L., & Birk, M. (2013). Deconstructing the touch experience. In *Proceedings of the 2013 acm international conference on interactive tabletops and surfaces* (pp. 199–208).

Zaharias, P., Michael, D., & Chrysanthou, Y. (2013). Learning through multi-touch interfaces in museum exhibits: an empirical investigation. *Educational Technology & Society*, *16*(3), 374–384.

Appendix

7.1 Appendix A - Pre-test questionnaire

A. Gender:
B. Age:
C. Previous experience with technology:

1	2	3	4	5	6	7
Not at all						Very much

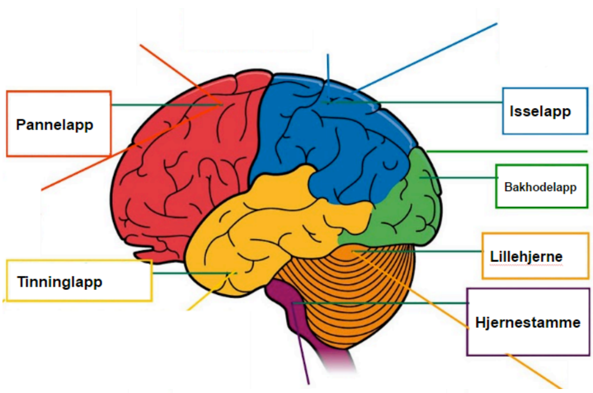
D. Previous experience with touch screens (e.g., big screens, ipads, mobile devices):

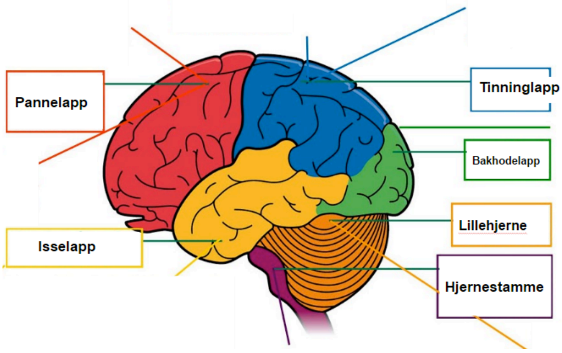
1	2	3	4	5	6	7
Not at all						Very much

E. Previous experience with big touch screens (bigger than a laptop size):

1	2	3	4	5	6	7
Not at all						Very much

1. Which image shows the correct placement of the parietal and temporal lobes?





2. More boys than girls have dyslexia

False, longitudinal research shows as many girls are affected as boys

True, this is linked with boys neurologically maturing later

3. Dyslexia affects __ % of the population and often occurs with ____

3-7%, ADHD

4. Which part of the brain is most closely associated with ADHD?

left-sided prefrontal cortex

right-sided prefrontal cortex

5. What is the cause of ADHD?

In a majority of cases, this is unfortunately unknown

The vast majority of cases have a causal link to population density

6. Which image shows the part of the brain affected by temporal lobe epilepsy?

Temporal lobe

7. What is the amygdala responsible for?

A. decision making

B. long-term memory

8. What the thalamus responsible for?

A. short term memory

B. sleep

9. Which image shows the thalamus?

Image A

Image B

10. Which part of the brain is responsible for language and attention?

A. **Cerebellum**

B. **Hippocampus**

7.2 Appendix B - Mid-test questionnaire and survey

Id:

1. Which gland produces insulin?

A. Pancreas

B. Pituitary

2. What are the common symptoms of dyslexia?

A. Trouble reading

B. Insomnia

3. Which brain functions are affected by ADHD?

A. the left-sided prefrontal cortex.

B. the right-sided prefrontal cortex.

4. Which image shows the part of the brain associated with ADHD?

5. What is the most common cause of temporal lobe epilepsy?

A. traumatic brain injury

B. headache

6. What treatments are effective for sufferers of temporal lobe epilepsy?

A. Antiepileptic drugs and surgical interventions

B. No treatment

7. Which region of the brain stores long term memory?

A **Hippocampus**

B **Thalamus**

8. What is the function of the hypothalamus

A **Body temperature**

B **Emotions**

9. Which part of the brain controls body temperature?

A **Hypothalamus**

B **Amygdala**

10. The cerebellum is responsible for ____

A **Language and attention**

B **Short time memory**

"Activity" here refers to the poster session

1. Please indicate how much you agree or disagree with the following statements regarding your satisfaction with the activity.

I am satisfied with the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I am pleased with the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

My decision to attend the activity was a wise one

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

2. Please indicate how much you agree or disagree with the following statements regarding your intention to use the activity in the future.

I intend to attend the activity in the future

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

My general intention to attend the activity in the future is very high

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I will regularly attend similar activities in the future

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I will think about attending similar activities

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

3. Please indicate how much you agree or disagree with the following statements regarding the easiness of the activity

The activity was easy

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I found the activity flexible

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

The process of the activity was clear and understandable

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

It was easy for me to attain skills with the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

4. Please indicate how much you agree or disagree with the following statements regarding your enjoyment with the activity.

Attending the activity was enjoyable

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Attending the activity was exciting

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I was feeling good in the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Attending the activity was boring

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

5. Please indicate how much you agree or disagree with the following statements regarding the usefulness of the activity:

I found the activity useful

	1	2	3	4	5	6	7	

Not at all ○ ○ ○ ○ ○ ○ ○ Very much

The activity improves my performance in biology and technologies.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

The activity enhances the effectiveness in biology and technologies.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

The activity increases my capabilities in biology and technologies

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

6. Please indicate how much control you believe you have over the activity:

I was able to follow the tasks of the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I have the knowledge and the ability to follow the tasks of the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

7. How you would characterize the activity:

Using this activity is compatible with most aspects of a digital archive/museum

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Using this activity fits my learning style

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Using this activity fits well with the way I like to engage in informal learning

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

8. How would you characterize your performance with the activity?

Being involved with the digital application I perform better than the acceptable level

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I perform better than can be expected from me

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I put in extra effort in my work

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I expend a great deal of effort carrying out my work

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I try to learn as more as possible

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application the quality of my learning is top-notch

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

9. Do you have any suggestions in order to improve the activity?

10. Here you can write whatever you want for the activity:

7.3 Appendix C - Post-test questionnaire, survey, Hedonic survey

Id:

1. What is the name of the hormone the testes produce?

A. Testosterone

B. Insuline

2. Which image shows the part of the brain associated with dyslexia?

A:

B:

(orange is considered highlighted)

3. What is the most common symptom of ADHD?

A. problems paying attention

B. abnormal aging

4. The temporal lobe is associated with temporal lobe epilepsy, which image points to the temporal lobe?

A:

B:

(orange is considered highlighted)

5. What is the most common symptom of temporal lobe epilepsy?

A. Lack of attention

B. Simple partial seizures

6. Which part of the brain is responsible for long-term memory?

A. Hypothalamus

B. Hippocampus

7. Which part of the brain regulates body temperature and hunger control?

A. Hypothalamus

B. Hippocampus Memory:

8. What is the function of the cerebellum?

A. Language and attention

B. Sleep and sensory interpretation

9. What is the function of the thalamus?

A. Language and attention

B. Sleep and sensory interpretation

10. What is the most common name for neurons?

A. Nerve Cells

B. Brain Cells

"Activity" here refers to the enhanced activity with the muti-touch application

1. Please indicate how much you agree or disagree with the following statements regarding your satisfaction with the activity.

I am satisfied with the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I am pleased with the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

My decision to attend the activity was a wise one

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

2. Please indicate how much you agree or disagree with the following statements regarding your intention to use the activity in the future.

I intend to attend the activity in the future

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

My general intention to attend the activity in the future is very high

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I will regularly attend similar activities in the future

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I will think about attending similar activities

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

3. Please indicate how much you agree or disagree with the following statements regarding the easiness of the activity

The activity was easy

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I found the activity flexible

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

The process of the activity was clear and understandable

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

It was easy for me to attain skills with the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

4. Please indicate how much you agree or disagree with the following statements regarding your enjoyment with the activity.

Attending the activity was enjoyable

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Attending the activity was exciting

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I was feeling good in the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Attending the activity was boring

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

5. Please indicate how much you agree or disagree with the following statements regarding the usefulness of the activity:

I found the activity useful

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

The activity improves my performance in biology and technologies.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

The activity enhances the effectiveness in biology and technologies.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

The activity increases my capabilities in biology and technologies

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

6. Please indicate how much control you believe you have over the activity:

I was able to follow the tasks of the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

I have the knowledge and the ability to follow the tasks of the activity

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

7. How you would characterize the activity:

Using this activity is compatible with most aspects of a digital archive/museum

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Using this activity fits my learning style

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Using this activity fits well with the way I like to engage in informal learning

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

8. How would you characterize your performance with the activity?

Being involved with the digital application I perform better than the acceptable level

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I perform better than can be expected from me

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I put in extra effort in my work

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I expend a great deal of effort carrying out my work

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application I try to learn as more as possible

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Being involved with the digital application the quality of my learning is top-notch

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

9. Do you have any suggestions in order to improve the activity?

10. Here you can write whatever you want for the activity:

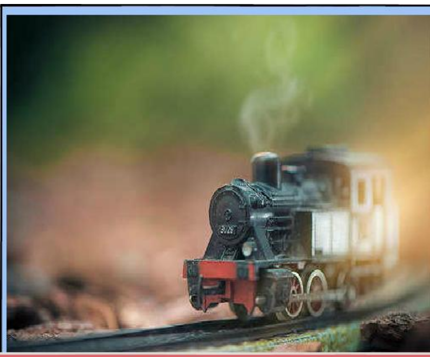
11. How would you characterize the experience you had with the multi-touch (MT) technology during the activity?

	Not at all Very much
MT is Effective	○ ○ ○ ○ ○ ○ ○
MT is Helpful	○ ○ ○ ○ ○ ○ ○
MT is Functional	○ ○ ○ ○ ○ ○ ○
MT is Practical	○ ○ ○ ○ ○ ○ ○
MT is Necessary	○ ○ ○ ○ ○ ○ ○

MT is Fun	○	○	○	○	○	○	○
MT is Exciting	○	○	○	○	○	○	○
MT is Delightful	○	○	○	○	○	○	○
MT is Enjoyable	○	○	○	○	○	○	○
MT is Thrilling	○	○	○	○	○	○	○
MT is Easy	○	○	○	○	○	○	○
MT is Flexible	○	○	○	○	○	○	○
MT is easy to attain skills	○	○	○	○	○	○	○
I will regularly use MT in the future	○	○	○	○	○	○	○
I intend to use MT in the future	○	○	○	○	○	○	○
I will think about using MT	○	○	○	○	○	○	○

7.4 Appendix D - Balsamiq prototype screenshots

Hvilken av disse to er en trikk?



Riktig!

0 Istid | 1 $ XP | 1 hint

lvl 1

Sprek bille

0 xp / 1000 xp

Feil...

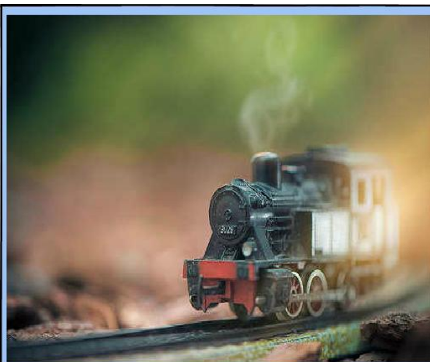
0 Istid | 1 $ XP | 1 hint

lvl 2

Blid spøkelse

200 xp / 1200 xp

Hvilken av disse to er en trikk?



Besvart

Angre?

Venstre

Høyre

45s

Tid gjenstående for dette spørsmålet

0 Istid | 1 $ XP | 1 hint

lvl 1

Sprek bille — 0 xp / 1000 xp

0 Istid | 1 $ XP | 1 hint

lvl 2

Blid spøkelse — 200 xp / 1200 xp

Hvilken av disse to er en trikk?

Venstre

Høyre

45s

Tid gjenstående for dette spørsmålet

0 Istid

1 $ XP

1 ? hint

lvl 1

Sprek bille

0 xp / 1000 xp

Venstre

Høyre

45s

Tid gjenstående for dette spørsmålet

0 Istid

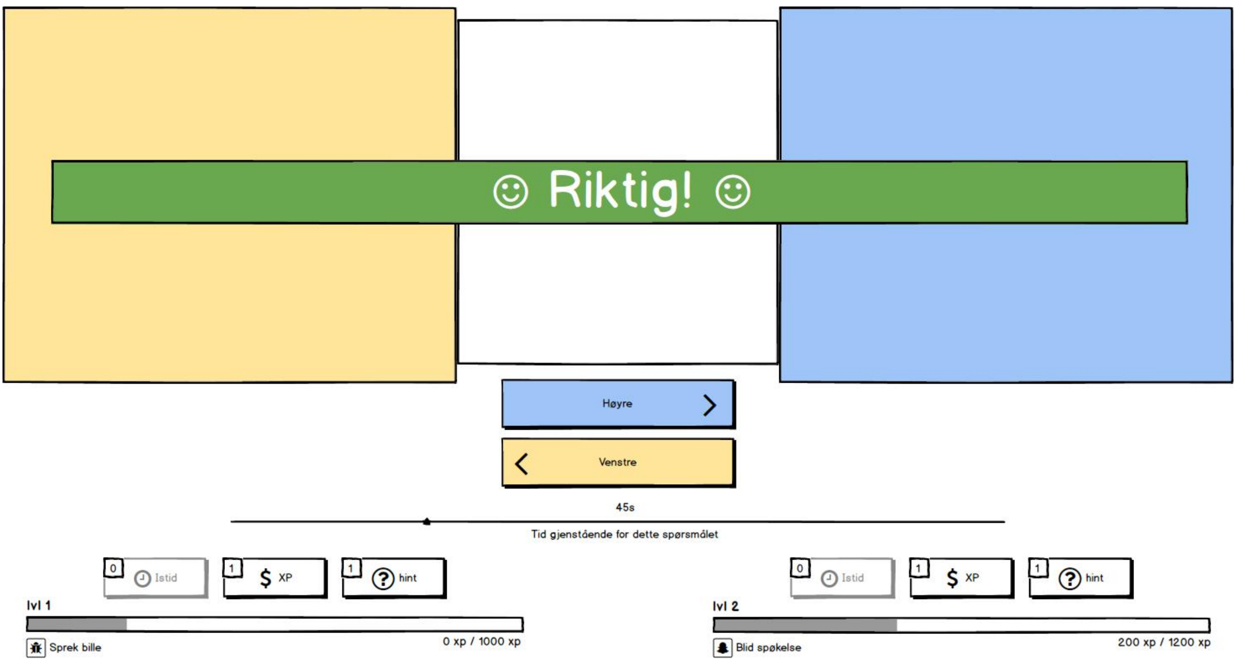
1 $ XP

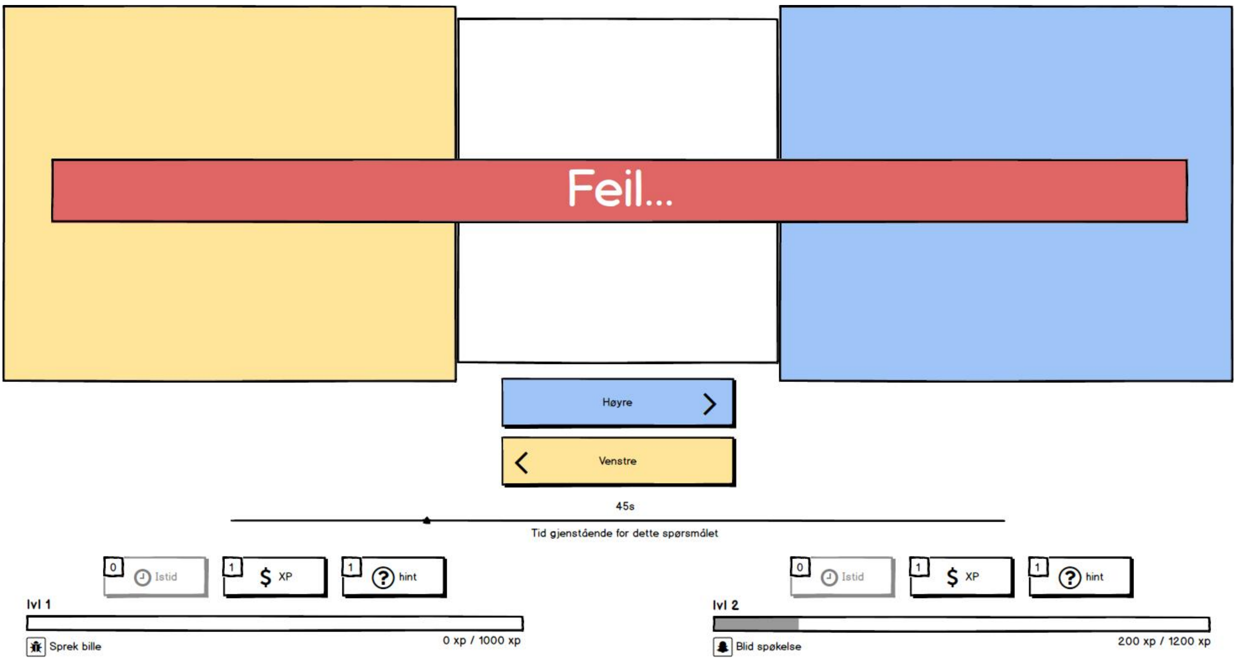
1 ? hint

lvl 2

Blid spøkelse

200 xp / 1200 xp


☺ Riktig! ☺
Høyre
Venstre
45s
Tid gjenstående for dette spørsmålet
0 Istid
1 $ XP
1 hint
lvl 1
Sprek bille
0 xp / 1000 xp
0 Istid
1 $ XP
1 hint
lvl 2
Blid spøkelse
200 xp / 1200 xp


Feil...
Høyre
Venstre
45s
Tid gjenstående for dette spørsmålet
0
Istid
1
XP
1
hint
lvl 1
Sprek bille
0 xp / 1000 xp
0
Istid
1
XP
1
hint
lvl 2
Blid spøkelse
200 xp / 1200 xp

Hvilken av disse to er en trikk?

Høyre

Venstre

45s

Tid gjenstående for dette spørsmålet

0 Istid | 1 $ XP | 1 hint

lvl 1

Sprek bille — 0 xp / 1000 xp

0 Istid | 1 $ XP | 1 hint

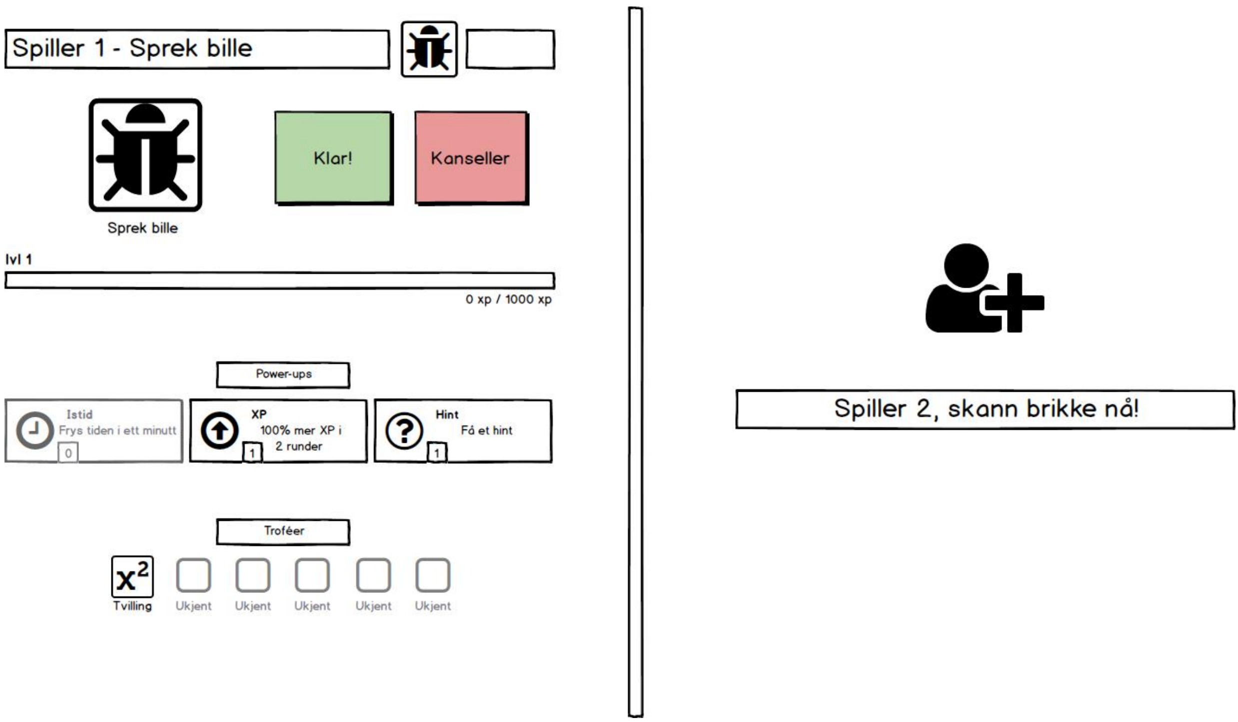
lvl 2

Blid spøkelse — 200 xp / 1200 xp

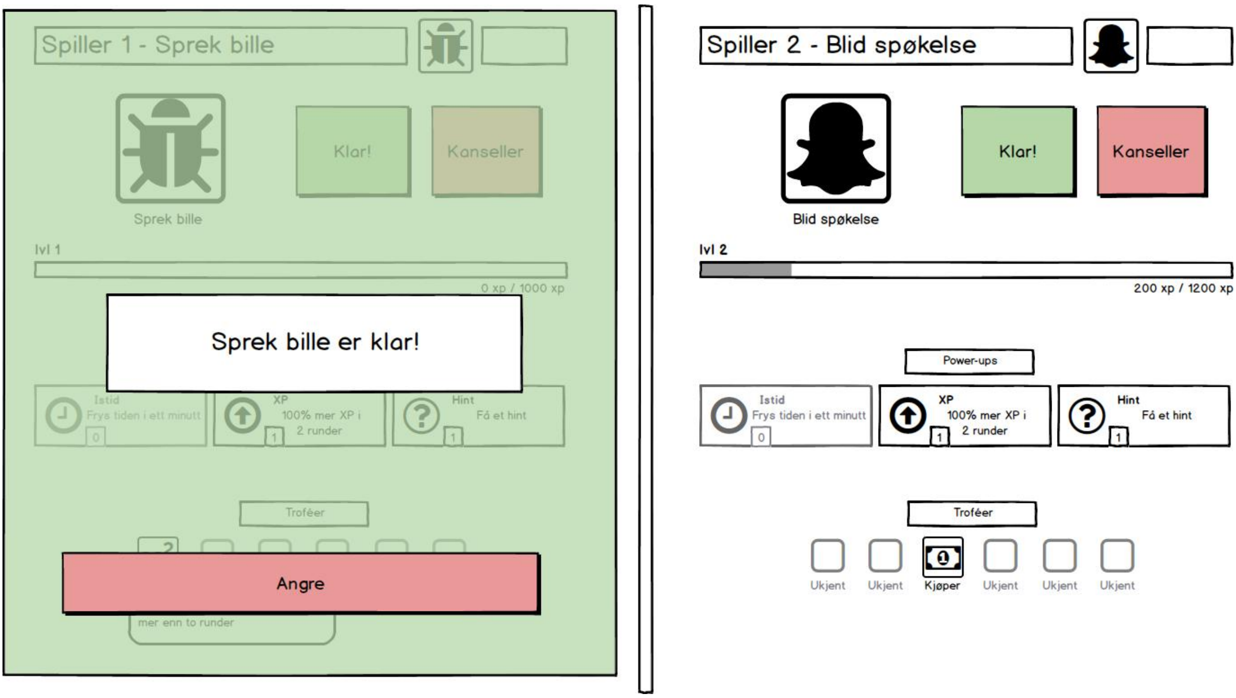
Velg ønsket spillemåte

Spille 1 mot 1!

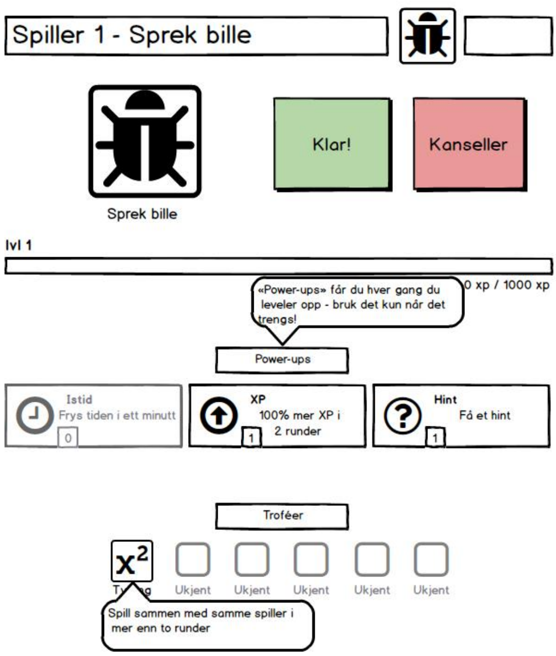
Spille sammen!



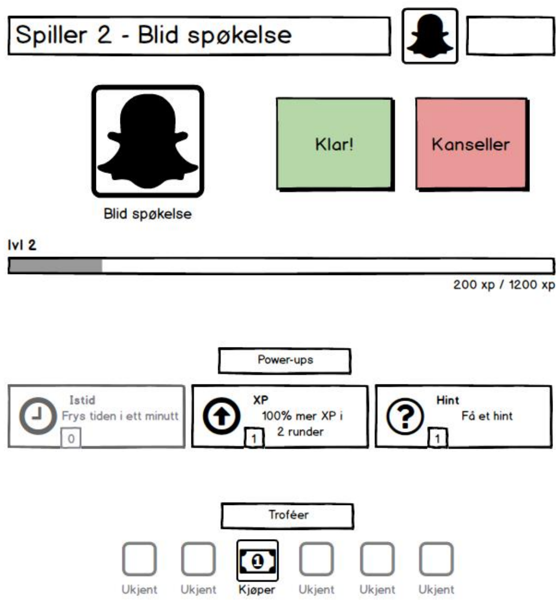
Spiller 1 - Sprek bille
Klar!
Kanseller
Sprek bille
lvl 1
0 xp / 1000 xp
Power-ups
Istid
Frys tiden i ett minutt
0
XP
100% mer XP i 2 runder
1
Hint
Få et hint
1
Troféer
Tvilling
Ukjent
Ukjent
Ukjent
Ukjent
Ukjent
Spiller 2, skann brikke nå!



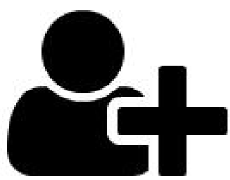
Spiller 1 - Sprek bille
Klar!
Kanseller
Sprek bille
lvl 1
0 xp / 1000 xp
Sprek bille er klar!
Istid
Frys tiden i ett minutt
0
XP
100% mer XP i 2 runder
1
Hint
Få et hint
1
Troféer
Angre
mer enn to runder
Spiller 2 - Blid spøkelse
Klar!
Kanseller
Blid spøkelse
lvl 2
200 xp / 1200 xp
Power-ups
Istid
Frys tiden i ett minutt
0
XP
100% mer XP i 2 runder
1
Hint
Få et hint
1
Troféer
Ukjent
Ukjent
Kjøper
Ukjent
Ukjent
Ukjent



Spiller 1 - Sprek bille
Klar!
Kanseller
Sprek bille
lvl 1
0 xp / 1000 xp
«Power-ups» får du hver gang du leveler opp - bruk det kun når det trengs!
Power-ups
Istid
Frys tiden i ett minutt
0
XP
100% mer XP i 2 runder
1
Hint
Få et hint
1
Troféer
Ukjent
Ukjent
Ukjent
Ukjent
Ukjent
Spill sammen med samme spiller i mer enn to runder



Spiller 2 - Blid spøkelse
Klar!
Kanseller
Blid spøkelse
lvl 2
200 xp / 1200 xp
Power-ups
Istid
Frys tiden i ett minutt
0
XP
100% mer XP i 2 runder
1
Hint
Få et hint
1
Troféer
Ukjent
Ukjent
Kjøper
Ukjent
Ukjent
Ukjent



Skann brikken din for å spille!

7.5 Appendix E - Likert-scale survey

Mid-test	AVG	SD
satisfaction_middle_average	5.101851852	1.349472429
intention_middle_average	4.486111111	1.404005042
easiness_middletest_average	5.270833333	0.9970939918
enjoyment_middletest_average	5.145833333	1.501636013
usefulness_middletest_average	4.604166667	1.302984212
control_middletest_average	5.625	1.203417752
activity_middletest_average	4.611111111	1.23056317
performance_middletest_average	4.20952381	0.5649669681

	Mid	**Post**
satisfaction	5.101851852	6.287037037
intention	4.486111111	5.284722222
easiness	5.270833333	5.576388889
enjoyment	5.145833333	6.236111111
usefulness	4.604166667	5.034722222
control	5.625	5.972222222
activity	4.611111111	5.611111111
performance	4.20952381	4.569444444

Post-test	AVG	SD
satisfaction_posttest_average	6.287037037	0.7731723058
intention_posttest_average	5.284722222	1.372110962
easiness_posttest_average	5.576388889	1.05698264
enjoyment_middletest_average	6.236111111	0.7767708902
usefulness_posttest_average	5.034722222	1.297184833
control_posttest_average	5.972222222	1.013793755
activity_posttest_average	5.611111111	0.8783100657
performance_posttest_average	4.569444444	0.5076791263

Diff	AVG	SD
satisfaction	1.185185185	-0.5763001236
intention	0.7986111111	1.372110962
easiness	0.3055555556	1.05698264
enjoyment	1.090277778	0.7767708902
usefulness	0.4305555556	1.297184833
control	0.3472222222	1.013793755
activity	1	0.8783100657
performance	0.3599206349	0.5076791263

	Mean (S.D.)		**T**	**P**
	Mid	**Post**		
Satisfaction	5.10 (1.35)	6.29 (0.77)	**4.895693168**	.000011

Intention	4.49 (1.40)	5.28 (1.37)	**3.718976813**	.000349	
Easiness	5.27 (1.00)	5.58 (1.06)	**2.017948311**	.025659	
Enjoyment	5.15 (1.50)	6.24 (0.78)	**4.128027464**	.000108	
Usefulness	4.60 (1.30)	5.03 (1.30)	**2.763893938**	.004525	
Control	5.63 (1.20)	5.97 (1.01)	**1.718977357**	.047234	
Activity	4.61 (1.23)	5.61 (0.88)	**4.810702354**	.000014	
Performance	4.21 (0.56)	4.58 (0.51)	**3.582452902**	.000512	

7.6 Appendix F - Shapiro and Wilcoxon calculations

pre_mid	mid_post	pre	mid	post
-0.08888888889	-0.1333333333	0.1111111111	0.2	0.3333333333
-0.3888888889	-0.3888888889	0.1111111111	0.5	0.8888888889
-0.2	-0.02222222222	0	0.2	0.2222222222
-0.3	-0.03333333333	0	0.3	0.3333333333
-0.1333333333	-0.08888888889	0.6666666667	0.8	0.8888888889
-0.3444444444	0.1222222222	0.5555555556	0.9	0.7777777778
-0.9	0.01111111111	0	0.9	0.8888888889
-0.9	0.01111111111	0	0.9	0.8888888889
-0.2222222222	0.2222222222	0.7777777778	1	0.7777777778
-1	0.1111111111	0	1	0.8888888889
-0.4666666667	-0.2	0.3333333333	0.8	1
-0.4777777778	0.03333333333	0.2222222222	0.7	0.6666666667
-0.5555555556	0.1111111111	0.4444444444	1	0.8888888889
-0.5555555556	0.1111111111	0.4444444444	1	0.8888888889
-0.2222222222	0.2222222222	0.7777777778	1	0.7777777778
-0.3444444444	0.01111111111	0.5555555556	0.9	0.8888888889
-0.5666666667	0.01111111111	0.3333333333	0.9	0.8888888889
-0.3888888889	0.1666666667	0.1111111111	0.5	0.3333333333
-0.4	0.2888888889	0	0.4	0.1111111111
-0.6	0.1555555556	0	0.6	0.4444444444
-0.3555555556	-0.08888888889	0.4444444444	0.8	0.8888888889
-0.6888888889	0.02222222222	0.1111111111	0.8	0.7777777778
-0.7777777778	0	0.2222222222	1	1
-0.7	0.1444444444	0	0.7	0.5555555556
-0.5888888889	-0.07777777778	0.1111111111	0.7	0.7777777778
0.06666666667	-0.1777777778	0.6666666667	0.6	0.7777777778
-0.3	-0.7	0	0.3	1
-0.5888888889	-0.1888888889	0.1111111111	0.7	0.8888888889
-0.2	-0.3555555556	0	0.2	0.5555555556
-0.3777777778	-0.06666666667	0.2222222222	0.6	0.6666666667
-0.4666666667	-0.08888888889	0.3333333333	0.8	0.8888888889
-0.2444444444	-0.08888888889	0.5555555556	0.8	0.8888888889
-0.6	0.04444444444	0	0.6	0.5555555556

-0.5777777778	0.02222222222	0.2222222222	0.8	0.7777777778
-0.4888888889	0.1555555556	0.1111111111	0.6	0.4444444444
-0.7	-0.07777777778	0	0.7	0.7777777778
Shapiro	Shapiro			
0.9394	0.003			
~normal distribution	**~reject normality**			
	Wilcoxon			
	0.934			
	~reject significance			